

Fri May 24 11:27:35 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: May 24, 2002, 10:12:57 ; Search time 101.42 Seconds
(without alignments)
516.928 Million cell updates/sec

Title: WALICK-934-95.PEP
Perfect score: 2530
Sequence: 1 MKPTSLWASAGALALIAAP.....GMOYVIALGGVSYSGSINS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_032802: *
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT: *
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT: *
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT: *
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT: *
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT: *
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT: *
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT: *
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT: *
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT: *
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT: *
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT: *
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT: *
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT: *
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT: *
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT: *
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT: *
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT: *
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT: *
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT: *
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: *
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT: *
23: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2478	97.9	579	19	AAW37873
2	2153	85.1	579	19	AAW37874
3	2145.3	84.8	579	19	AAW37876
4	2065.5	81.6	578	21	AAW37875
5	413	16.3	754	21	AAW35987
6	408	16.1	742	11	AAW05235
7	407.5	16.1	738	13	AAW02192
8	406.5	16.1	738	12	AAW13993
9	274.5	10.8	740	20	AAW95019
10	137	5.4	443	22	AAW24430
11	122.5	4.8	715	22	AAW61738

12	122.5	4.8	948	22	ABG25909
13	122.5	4.8	1510	22	ABG21573
14	122.5	4.8	910	22	AAW38007
15	121.5	4.7	886	22	AAW32914
16	119	4.7	886	22	AAW33996
17	118.5	4.7	770	22	AAW32701
18	118	4.7	770	22	AAW30824
19	115.5	4.6	501	18	AAW10645
20	114.5	4.5	501	18	AAW10644
21	113.5	4.5	1016	22	AAW33449
22	113	4.5	824	18	AAW09614
23	111.5	4.4	501	18	AAW26108
24	111.5	4.4	501	18	AAW26109
25	110.5	4.4	501	18	AAW26107
26	110.5	4.4	501	21	AAW99981
27	110.5	4.4	594	12	AAW10005
28	108.5	4.3	530	22	AAW32080
29	108.5	4.3	530	22	AAW12236
30	108	4.3	500	18	AAW24582
31	107.5	4.2	501	18	AAW32082
32	107.5	4.2	501	18	AAW15110
33	107.5	4.2	501	18	AAW26113
34	107.5	4.2	501	18	AAW26115
35	107.5	4.2	501	18	AAW26116
36	107.5	4.2	269	11	AAW26106
37	107	4.2	269	15	AAW53295
38	107	4.2	269	15	AAW82772
39	106.5	4.2	404	22	AAW82772
40	106.5	4.2	501	18	AAW15109
41	106.5	4.2	501	18	AAW15111
42	106.5	4.2	501	18	AAW10646
43	106.5	4.2	593	11	AAW07014
44	106.5	4.2	2599	21	AAW75098
45	106	4.2	269	10	AAW94788

ALIGNMENTS

RESULT 1	AAW37873	standard; Protein: 579 AA.
ID	AAW37873	
AC	AAW37873	
XX	10-AUG-1998	(first entry)
DT	XX	
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulononic acid; L-ascorbic; inhibition.	
XX	XX	
OS	Gluconobacter oxidans.	
XX	XX	
FX	Key	Location/Qualifiers
FT	Peptide	1..23 "signal peptide"
FT	Protein	24..579 "mature protein"
FT	Protein	/note="mature protein"
XX	XX	
PN	EP832974-A2.	
PD	01-APR-1998.	
XX	XX	
PF	11-SEP-1997;	97EP-0115801.
XX	XX	
PR	19-SEP-1996;	96EP-0115001.
XX	XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX	Asakura A, Hoshino T, Ojima S, Shiojoh M, Tomiyama N;	
PI	Novel human diageno	
XX	S. epidermidis ope	
XX	S. epidermidis sequ	
XX	Human protein sequ	
XX	S. epidermidis ope	
XX	Amino acid sequenc	
XX	HPV6 mutant L1 pro	
XX	HPV6 mutant L1 pro	
XX	Staphylococcus aur	
XX	Pathogenic Staphyl	
XX	HPV11 mutant L1 pr	
XX	HPV11 mutant L1 pr	
XX	HPV11 mutant L1 pr	
XX	HPV 6 L1 protein m	
XX	Streptococcus GX78	
XX	Human HPV6 L1 prot	
XX	Human PRO3440 poly	
XX	HPV6 L1 protein.	
XX	Human HPV6 L1 prot	
XX	Human HPV6 L1 prot	
XX	HPV11 mutant L1 pr	
XX	HPV11 mutant L1 pr	
XX	HPV11 mutant L1 pr	
XX	Protein G variant.	
XX	IgG-binding Strept	
XX	S. epidermidis ope	
XX	Human HPV6 L1 prot	
XX	Human HPV6 L1 prot	
XX	HPV11 mutant L1 pr	
XX	Protein G variant	
XX	Neisseria meningit	
XX	Protein G variant.	

DR WPI: 1998-195228/18.
 N-PSDB: AAV29051.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 35-37; 59pp; English.
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX Sequence 579 AA;
 SO

Query Match 97.9%; Score 2478; DB 19; Length 579;
 Best Local Similarity 84.9%; Pred. No. 2.2e-210;
 Matches 472; Conservative 0; Mismatches 0; Indels 84; Gaps 1;
 OY 1 MKPTSLMASGALALLAARAFQVTPYDELANPPAGEWISYGOENYRHSPLTOIT 60
 DB 1 mkptslmasgalaallaaraafqvtptdelanppagewisygngenyhspltlgt 60
 OY 61 TENVGOLQLVWARGMPGKQVTPPLIHGVMYLAN----- 95
 DB 61 tenvgqlqlvwarmpgkqvtpplihgvmylan----- 95
 OY 96 ----- 96
 DB 121 latlnsfgeptirmalylgtlnvfyvswdnhlvaldtaqvtfdvdrqggedmvsnsqpl 180
 OY 97 VANGYIVAGSTCYSPFCFVSGHDSATGELMRNYFIRPACBEGDETNWGNDEYARMNTG 156
 DB 181 vangyivagstcyspfcfvsghdsatgelmrnyfirpaegegetwngndeyearmmtg 156
 OY 157 AMGOITDPVTNLVHYGSTAVGPASSTQRTGRTGTLGNTNRFAVRPDGCEIYWRHQTLP 216
 DB 241 awgiltypvtlnlvhygstavgpasstqrtgtrtltfavrpdgceiylwrhqtlp 216
 OY 217 RDNNDDECTEMAVTNVDQVSTEMEGLOSINPMAAGERRVLTGVPCKTGTMOFDAET 276
 DB 301 rdnnddectemavtnvdqvstemeqlgsinpmaagerlvltgvpcktgtmofdaet 276
 OY 277 GEPLWARDNTNOMTIESIDENGIYTVNEDAILKELDEVYDVPFLGGRDMPSAALNPD 336
 DB 361 geplwardntnontiesidenyitvnedailkeldevydvptflggrdmpsaalnps 336
 OY 337 GIYFPLNNVCYDMMAVDQERTSHDYNTSNVTKLPPGKDMIGRDAIDISTGRTLMSVE 396
 DB 421 giyfplnnvcydmmaavdqertshdyntsnvtklppgkdmigrdaidistgtrtlmsve 396
 OY 397 RAAANYSVIVSTGGVLENGTDTRYPRALSOETGTTLMQTRLATVASGOAISTEVGMOMY 480
 DB 481 raaanysvivistggvlengtdtrypralsetgttltmqtrlatvasgoaistevgmomy 480
 OY 457 VALAGGVSTGSGTNS 472
 DB 541 valaggvstsgtlns 556

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein; 579 AA.
 XX
 AC AAM37874;
 XX

DT 10-AUG-1998 (first entry)
 XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 DE
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 XX 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note="signal peptide"
 FT /note="mature protein"
 PN EF832974-A2.
 PD 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 PF 19-SEP-1996; 96EP-0115001.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 38-40; 59pp; English.
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX Sequence 579 AA;
 SO
 Query Match 85.1%; Score 2153; DB 19; Length 579;
 Best Local Similarity 72.8%; Pred. No. 1.3e-181;
 Matches 402; Conservative 32; Mismatches 34; Indels 84; Gaps 1;
 OY 1 MKPTSLMASGALALLAARAFQVTPYDELANPPAGEWISYGOENYRHSPLTOIT 60
 DB 1 mkptslmasgalaallaaraafqvtptdelanppagewisygngenyhspltlgt 60
 OY 61 TENVGOLQLVWARGMPGKQVTPPLIHGVMYLAN----- 95
 DB 61 tenvgqlqlvwarmpgkqvtpplihgvmylan----- 95
 OY 96 ----- 96
 DB 121 latlnsfgeptirmalylgtlnvfyvswdnhlvaldtaqvtfdvdrqggedmvsnsqpl 180
 OY 97 VANGYIVAGSTCYSPFCFVSGHDSATGELMRNYFIRPACBEGDETNWGNDEYARMNTG 156
 DB 181 vangyivagstcyspfcfvsghdsatgelmrnyfirpaegegetwngndeyearmmtg 156
 OY 157 AMGOITDPVTNLVHYGSTAVGPASSTQRTGRTGTLGNTNRFAVRPDGCEIYWRHQTLP 216
 DB 241 awgiltypvtlnlvhygstavgpasstqrtgtrtltfavrpdgceiylwrhqtlp 216


```

OY 217 RDNDDECTFEMAVTVNDVQSTENEGLOSINPNAATGERRVLGVPCKTGMQFPAET 276
DB 301 rdnddectfemavtvndvqstendeglosinpnatgerrvlgvpcktgmlwqfdaet 360
OY 277 GEELMARDTQNMIESIDENGIVTNEEDALIKELDEYDVCPTFLGGRDMPALNPDS 336
DB 361 geflwardtqnmiesidengivtnedalikeldydvcpfllggrdmpaalnpds 420
OY 337 GIVFIPLNVCYDMAVDOEFTSMQVNTSVNTKLPCKDMIGRIDADISTGRTLSVE 396
DB 421 givfiplnvcymavdoeftsmqvntsvntklpckdmigrdadistgtrtlswae 480
OY 397 RAANSPVLSSTGGVTFNGCTDRYFRALSOETGELTMOTRLATVASGOAISYVDQMOY 456
DB 481 reasnspvlsstggvtfngctdryfralsetgelmotrlatvasgoaisyvdmoy 540
OY 457 VAIAGGVSYGS 468
DB 541 vaiaggvsgtys 552

```

RESULT 3

AAW37876 standard; Protein: 579 AA.

AAW37876;

10-ANG-1998 (first entry)

Alcohol and/or aldehyde dehydrogenase B amino acid sequence.

Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulononic acid; L-ascorbic; inhibition.

Glucobacter oxydans.

Location/Qualifiers

Key 1..23

Peptide /note="signal peptide"

Protein 24..579

/note="mature protein"

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB; AAW29054.

Recombinant Glucobacter oxydans alcohol and/or aldehyde.

dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.

L-sorbose or D-sorbitol to 2-keto-L-gulononic acid

Claim 1; Pages 44-46; 59pp; English.

This is the amino acid sequence for the Glucobacter oxydans

alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes

or recombinant organisms can be used to convert suitable substrates

to aldehydes, ketones or carboxylic acids, especially to convert

L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be

converted to L-ascorbic acid by standard procedures. The derivatives

of ADH enzymes have desired substrate specificity, higher affinity

to a substrate, lower affinity to an inhibitory compound, higher

CC stability against temperature and/or pH and higher catalytic speed.

XX Sequence 579 AA;

84.8%; Score 2145.5; DB 19; Length 579;

Best Local Similarity 71.6%; Pred. No. 5.8e-181; Indels 85; Gaps 1;

Matches 399; Conservative 34; Mismatches 39;

1 MKPTSLMWSACALALLAAPFAOYPTVTDLAPNAPGEMISYGOENRHSPLTQIT 60

1 mpftllrtaavvllltpaafayvpltdellapnpgewinyngenytrhspltlqit 60

61 TENVGOLQIYMAKMPCKVQYTPLIHDGVMYLAN----- 95

61 adnvgqlqlywalgmaagavqytpmldgvmylanpvgdylqaldaqtlwehrrqlpa 120

96 ----- 95

121 valnagdkrkrgvalyatslyfsswdnhllaldmetgvldevrgsgedjlsntltp 180

96 IVANGVIVAGSTCOYSPFGCFVSGHDSATGEELMKNYIFPRAGEGDETWGNDYEAHMT 155

181 Ivaugvivaagstcgyfsgcfvsgshdsatgeelmknhifprageegdetwgndearhmt 240

156 GAMQITVDYPTNLYHGSTAVGPASETORGPCTGTLGNTFRVAPDGTGEIYWRQTL 215

241 gvwgqitldyptnlvlyhgstavgpasetorgpctgtlgntrfvapdgtgeiylwrqtl 300

216 PRDNDDECTFEMAVTVNDVQSTENEGLOSINPNAATGERRVLGVPCKTGMQFPAE 275

301 prdnddectfemavtvndvqstendeglosinpnatgerrvlgvpcktgmlwqfdaa 360

276 TGEELMARDTQNMIESIDENGIVTNEEDALIKELDEYDVCPTFLGGRDMPALNPDS 335

361 tgeelmardtqnmiesidengivtnedalikeldydvcpfllggrdmpaalnpds 420

421 tglfiplnvcymavdoeftsmqvntsvntklpckdmigrdadistgtrtlswae 480

336 SGIVFIPLNVCYDMAVDOEFTSMQVNTSVNTKLPCKDMIGRIDADISTGRTLSV 395

481 sgivfiplnvcymavdoeftsmqvntsvntklpckdmigrdadistgtrtlswae 540

396 ERAANSPVLSSTGGVTFNGCTDRYFRALSOETGELTMOTRLATVASGOAISYVDQMOY 456

481 erasnspvlsstggvtfngctdryfralsetgelmotrlatvasgoaisyvdmoy 540

456 YVAIAGGVSYGSGLNS 472

541 yvaiaggvsgtysglns 557

RESULT 4

AAW37875 standard; Protein: 578 AA.

AAW37875;

10-ANG-1998 (first entry)

Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulononic acid; L-ascorbic; inhibition.

Glucobacter oxydans.

Location/Qualifiers

Key 1..23

Peptide /note="signal peptide"

Protein 24..578

/note="mature protein"

EP832974-A2.

XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 PF 19-SEP-1996; 96EP-0115001.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 PI WPI: 1996-195228/18.
 DR N-PSDB: AAV29053.
 DR
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1; Pages 41-43; 59pp; English.
 CC This is the amino acid sequence for the gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of AADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 578 AA;

Query Match
 Best Local Similarity 81.6%; Score 2065.5; DB 19; Length 578;
 Matches 388; Conservative 37; Mismatches 45; Indels 85; Gaps 2;

QY 1 MKPTSLMASGALALALAPAPAOVTPYDELANPAGEMISYGOQENYRHSPLTOIT 60
 DB 1 mktllqgsaalivgtipalaq-taltdemlanpagewinygqenyrhsplttgt 59
 QY 61 TENVGLOLVWARGKOPKQVQVTPPLIHGVNMYLAN----- 95
 DB 60 adhvqqlqlvwarigmeagkivtprlvhgvmylanpddvgaialdaatgdlwehrrqjpn 119
 QY 96 ----- 1
 DB 120 iatlnsfgeptrgmalvgtlnyfvswdnhvaldstlqgvfvdvdrdggtldmvsnsqpi 179
 QY 97 VANGYIVAGSVCQYSPFCGCVSGHDSATGELMRNYFIRAGBEDGMNDYKARMTG 156
 DB 180 vanyivagstcqvspfgcfvsgshdsatgelmrnfiraggegetwvndykarwmtg 239
 QY 157 AMGOITVDYPTNLVHGSTVAPASFOGCTREGTLKGNTRFARVPRPTGIVWRHOLP 216
 DB 240 twgqitvprvgvlvhygstvapaeeqfgrlvgsmytlntrfavreptgslvwrhqlp 299
 QY 217 RDNDDOCTEEMWNTVNDVOPSTEMEGLOSINPNATGERRVLGVCKTGMQFPAET 276
 DB 300 rdnddgecttemwvndvqpsaameglhainpdaatgerrvvgvpcnkgtlmvqfaet 359
 QY 277 GEFLIARDNTYQNMIESIDENGIVTVDALIKELDVEYDVCPTEFLGGRDWPSSAALNPPS 336
 DB 360 geflwardtsgnltvsgvpgdlvhnmedlvvtelevayelcplflggridwpsaalnpp 419
 QY 337 GIEFELNVCYDMAVDEFTSMOVNTSNVTKLPKGMIGRIDAIDISGRITLWSE 396
 DB 420 gylfelflmacsgmclavqefssldvynvslkpsgsnmgridaidsgrtlwseae 479
 QY 397 RAAATSPVLSTGGVLEFNGGTRDRFRALSGENGETIMQRLATVAVSGAISYEVDGMOY 456
 DB 480 ryasnayplstggvlefnngtldrfralsgecgetlwtqrlatvavsgaisyeldgvy 539

QY 457 VALAGGVSYSGSLN 471
 DB 540 valgrggtsgysnln 554

RESULT 5
 ID AAB35987 standard; Protein; 754 AA.
 XX AAB35987;
 AC AAB35987;
 DT 01-MAR-2001 (first entry)

DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulonic acid.

OS Gluconobacter oxydans.
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.

PF 23-APR-1999; 99MO-IB00736.
 XX
 PR 23-APR-1999; 99MO-IB00736.

PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.

PI Choi E, Rhee S, Lee E;
 XX
 DR WPI: 2000-687351/67.
 XX
 DR N-PSDB: AAC83153.

PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol
 XX
 PS Claim 1; Fig 8; 96pp; English.

CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulonic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 XX
 XX Sequence 754 AA;

Query Match
 Best Local Similarity 16.3%; Score 413; DB 21; Length 754;
 Matches 143; Conservative 72; Mismatches 222; Indels 164; Gaps 15;

QY 14 LALLAAPFAOVTPV-----TDELANPAGEMISYGOQENYRHSPLTOITENYQ 66
 DB 17 lgcnaalafcatlspvalaedgtalnadhpqdmmsygrtysqgrysplqdkdhaan 76
 QY 67 IQLVWARGMQGKVOY-TPLIHGVNMYL-----ANIVAN 99
 DB 77 lklawhyldlmgsgqetplivdgymyaltwsmkaldatgklwskydpkypgnladr 136
 QY 100 G----- 100

Db 137 gcdltvnrqaaaymgkvfygftdgrllaladaetgklywsvlytpkcaqjghnqstvtvda 196
QY 101 -----VIVAGSTCOYSPFCGFSGHSATGELMNRNFIIPRAGECD----- 142
Db 197 prlkgkvllgnggaeafgarg-fvlaydaetgkmwrtffvcpndkppdaaadvlmek 255
QY 143 --ETWGGNDYARHMTG---AMGQITTDPTNLVHYGSTAVGPASEIRGRPGTILGTNT 197
Db 256 aypwagkgagwkgggggtwgdsllydptdlylyvgngspwkyktrsegkgnlflgs 315
QY 198 RFVAVRPTGELIWMHQTLPDNDQECTFEEMAVTNVDVOPSTEMEGLOSTINPNATGE-R 256
Db 316 lvalnptgkyvwhfqtgtpmdqwdytsvqimalampv-----ngemt 358
QY 257 RVLTVGPKTGTMMQOPAEETGEELMARDNYONMIESIDE-NGIYVNEDELKELVEY 315
Db 359 hvlvnap-kngfflyldakgkfgskpytyenwangldpvtgtrpnydpdalwtlmgkpw 417
QY 316 DWCPTEFGGRDWPSSALNPDSCGIFIFLNNVCY-----DMNAVDEETSMGV 362
Db 418 yslpgqlgghmfaamayspckllylpaqvyrfvydpkgkfahdswnljldmk191 477
QY 363 YNNSN---VTKLPKGMIGRIDALIDISTGRILMSVERAANYSPLSTGGVLEFNGCT 418
Db 478 lddnpgkhdakqfllkdklyvawdpqkaafvdkhgwnglalatagvllfgla 537
QY 419 DRYFRALSOETGELMOTRLATVASGOAISYEVDGMQYVAI---AG-----GGVSYGS 468
Db 538 ngethaydattgkdlitfpaqsalppvtylangkyvavevgwgllypfllygvaarts 597
QY 469 G 469
Db 598 g 598

RESULT 6

AAR05235 standard; protein: 742 AA.
ID AAR05235;
AC AAR05235;
XX 04-AUG-1990 (first entry)
XX DT Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
XX DE 1-2229 of alcohol dehydrogenase (ADH) gene.
XX KW Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
XX KM Acetobacter aceti K1006 (FERM-7528); plasmid PAA721.
XX OS Acetobacter aceti strain K1006 (FERM-7528).
XX PN JP02000452-A.
XX PD 05-JAN-1990.
XX PF 30-OCT-1987; 87JP-0273190.
XX PR 30-OCT-1987; 87JP-0273190, JP-075069.
XX PA (QPP) QP CORP (QPPJ-).
XX DR WPI: 1990-047990/07.
XX DR N-PSDB; AAG91811.
XX PT Cloning DNA, plasmid and microbe contg. it
XX PT contains alcohol dehydrogenase gene derived from Acetobacter
XX PT aceti K1006 strain
XX PS Disclosure: Fig 4; 8pp; Japanese.
XX CC Also new are a recombinant plasmid contg. its encoding DNA, and a
XX CC microorganism transformed with the plasmid.

CC A DNA fragment was obt'd. from A. aceti
CC strain K1006 and was transferred to an ADH-defective strain. The
CC resulting plasmid, PAA721, contg. its gene was inserted into A.
CC pasteurianus strain IFO 3191 by the triparent method. Its gene is
CC useful for improving Acetobacter culturing.
SQ Sequence 742 AA:

Query Match 16.1%; Score 408; DB 11; Length 742;
Best Local Similarity 24.2%; Pred. No. 4,3e-27;
Matches 149; Conservative 78; Mismatches 192; Indels 196; Gaps 22;

QY 11 AGALALLAFAFAQVRYPTDEL--ANPPAGEWISYQONENYHSPLOTITENYGOLO 68
Db 23 aalpyaavpaagaqngtgealhhadhpewlsygrtseqlyspldqinsvnygk 82
QY 69 LVMARGHQPCKVQ-VTPLIHDGVNLT-----ANI----- 96
Db 83 llyytlidnrgaetrlvvvgimyatlmwskmealdaetgkllwqyqpkvgnladdgc 142
QY 97 ----- 96
Db 143 cdtvnrqaaaymgkvfygftdgrllvaadaktgkxwavnlpadaslgkqrsytdgavr 202
QY 97 VANGVIVAGSTCOYSPFCG--FVSGHDSATGELMNRNFIIPRAGECD----- 142
Db 203 vakglvllgn--ggaefgargfvaafdaetgkllkvtlyvnpknepdhaasdlmkka 260
QY 143 --ETWGGNDYARHMTG---GAMQITTDPTNLVHYGSTAVGPASEIRGRPGTILY--- 193
Db 261 ykewgp-kgaawrggggtwgdsllydpsdlly---lavg-----ngspwkykyise 309
QY 194 --GTN---TRFVVRPTGELIWMHQTLPDNDQECTFEEMAVTNVDVOPSTEMEGLOSI 247
Db 310 gisnllglslvalkpegeywhfgaepmdqwdytsvqimclampvk-----EN 297
QY 248 NPNATGE-RRVLTGVPCKTGTMMQOPDAETGEELMARDNYONMIESID----- 358
Db 359 -----gemrhlvnap-kngfflyldakgkfgskpytyenwangldpvtgtrpnydp 411
QY 298 GIYVNEDELKELVEYDYCPTEFGGRDWPSSALNPDSCGIFIFLNNVCY----- 348
Db 412 gilyclng-----kfwygl-pgplgahmfaamayspckllylpaqvyrfvydpkgkfahdswnljldmk191 477
QY 349 -----DMNAVDEETSMGVYNTSVTKLPKGMIGRIDALIDISTGRILMSVERAANYSPLSTGGVLEFNGCT 418
Db 464 kphadswngvlgmtknglpdpce-aitayikdlhgwllawdpvkmelwvkhkxgpnng 522
QY 405 VSTGGVLEFNGCTDRYFRALSOETGELMOTRLATVASGOAISYEVDGMQYVAI---AG-----GGVSYGS 468
Db 523 llatgdlilfglangefhaydatngsdlkyfdaqsqliapmtlysvngkyvavevgw 582
QY 459 ----IAGGVSYGS 469
Db 583 gilylmngvgrtsg 597

RESULT 7

AAR20192 standard; protein: 738 AA.
ID AAR20192;
AC AAR20192;
XX 16-APR-1992 (first entry)
XX DT ADH complex protein (mol.wt. 72.000).
XX DE Alcohol dehydrogenase; acetic acid; fermentation.
XX KM Acetobacter altoacetigenes NR-24.
XX OS JP03266988-A.
XX PN


```

OY 143 -----ETWGNDEYKARMTGA-----WGQITDPVTNLVHGSTAVGASFOR 185
XX          |||          |||          |||          |||          |||
DB 255 vlmnkayqlws-----pfgawlrqggggtvdwdsivypvadlvlylgvngspwnykyr 307
OY 186 GTPGGLVSTNRFARPDTEGEIWRHQTLPRDNDQECTFEKAMVTNVQOSTEMGLQ 245
XX          |||          |||          |||          |||          |||
DB 308 segkgnllgslvalkpetgeywhfgeipmqdwtfsdqqlmldlpr----- 357
OY 246 SIMPNAATGERRVLGVCCKGTMMQPDATGEFLMARDNTYONMIESID-ENGIVYVNF-304
XX          |||          |||          |||          |||          |||
DB 358 -----ngethivlvharknngfyildantgetisgknyvvnwasgldpktrpiynp 410
OY 305 DALIKELDEYDVCPTFLGGRDMPALNPDGTYTPIPLNNCYDMMAVDQECT-SMDVY 363
XX          |||          |||          |||          |||          |||
DB 411 dalytltygkewyigpdlgghnaaatspkrtylylpaqgvplyltnvggftlprdsaw 470
OY 364 NTS---NVTKLPRG-----KDMIGRIDALIDSTGRTLMSVERAANYSPVLSTGGV 412
XX          |||          |||          |||          |||          |||
DB 471 nlglamkvqipdspeakqafvkdikgylvawdpqkgaearvhkqpmngllatqgd 530
OY 413 LFNSTDRYFRALSOETGETLMQRLATVAGSQAISYEVDMQVVAIAGG 462
XX          |||          |||          |||          |||          |||
DB 531 lfgqlangefhayatngsdlfnfaadsgllapvlylangkqyvaevg 580

RESULT 9
AAW95019
ID AAW95019 standard; Protein: 740 AA.
XX
AC AAW95019;
XX
DT 21-MAY-1999 (first entry)
XX
DE Sorbitol dehydrogenase (SDH) protein sequence.
XX
XX Sorbitol dehydrogenase; SDH; open reading frame; ORF2; L-sorbose;
KM D-sorbitol; vitamin C; enzyme.
XX
XX Glucanobacter suboxydans.
OS
XX
FH Key location/Qualifiers
FT Peptide 1..24
FT /note= "signal sequence"
FT /note= "mature protein"
FT Protein 25..740
XX
XX EP897984-A2.
XX
PD 24-FEB-1999.
XX
PF 13-AUG-1998; 98EP-0115231.
XX
PR 21-AUG-1997; 97EP-0114432.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N.
XX
XX WPI; 1999-134646/12.
XX
XX N-PSDB; AAX21501.
XX
XX New D-sorbitol dehydrogenase gene and recombinant protein - useful
PT for production of L-sorbose, an intermediate in vitamin C production
XX
XX Claim 1; Fig 3A-D; 39pp; English.
XX
XX This represents a sorbitol dehydrogenase (SDH) protein. The DNA
XX encoding the SDH enzyme also encodes an open reading frame (ORF2)
XX product upstream of the SDH open reading frame, needed for SDH activity
XX in vivo. Host cells transformed by a vector comprising the SDH DNA
XX sequence are used for the recombinant expression of the sorbitol
XX dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
XX producing L-sorbose from D-sorbitol. L-sorbose is an important

```

```

CC intermediate in vitamin C production.
XX
XX Sequence 740 AA;
SQ
Query Match 10.8%; Score 274.5; DB 20; Length 740;
Best Local Similarity 20.6%; Pred. No. 2.8e-15;
Matches 143; Conservative 70; Mismatches 167; Indels 313; Gaps 33;

OY 23 AOVTEVDTELLANPAGEMISYONGNENRHSPLQITTEENVGQQL--VWARKH--QPG 78
XX          |||          |||          |||          |||          |||
DB 86 sqvmapaqgsanpargdwaaygriddhqtlyspiselpenssklkvaftvlyhtgysp;pg 145
OY 79 KV-----QVTPV-IHDGYVLANIVANGVIVAGSTCOYSPFGCFVSGHSAATGEELMR-- 130
XX          |||          |||          |||          |||          |||
DB 146 qynkwaatetpikvdyg-lycsam-ndilk1-----dpagkqkwrtn 187
OY 131 -----NYF-----IPRA----- 137
XX          |||          |||          |||          |||          |||
DB 188 vdvkynslpylaackgvlyftsvvpeqgpcmrlliegtldmrliavdaetqdfcpnfigh 247
OY 138 -----GE-----EGDETM-----GNDYEARM 154
XX          |||          |||          |||          |||          |||
DB 248 gqgvnlmgjgesvpgfvmptappvngvvvnhvldgqtrwapsvlyrydaesgkf 307
OY 155 TGAN-----GQITDPVTNLVHGSTAVGASFORQRTPGTLYGTNR----- 198
XX          |||          |||          |||          |||          |||
DB 308 vwadvnmsgrspay-rylvlaaverltpjldtrrg-----gsrlprdnasady 359
OY 199 -FAVRPD-----TGEIWRHQTLPRDNDQECTFEKAMVTNVQOSTEME 242
XX          |||          |||          |||          |||          |||
DB 360 ysalsdaenkvsavvaivdksgsprwvfgqahdvwdydgsgatl-----ndmp 411
OY 243 GIGSINPNAATGERRVLGVCCKGTMMQPDATGEFLMARDNTYONMIESIDEN----- 297
XX          |||          |||          |||          |||          |||
DB 412 g-----pdqgtvpallm---pkrgqtfvldrttgk-----pdlpveerpasp 452
OY 298 GIV-----TVNEDAI-----LKELD----- 312
XX          |||          |||          |||          |||          |||
DB 453 gvlpgdprspcpwsgvmpalrtpdlketdmwgmsspidqfckfriranygetfpssv 512
OY 313 ----VEYDVCPTFLGGRDMPALNPDGTYTPIPLN-NVCYDMMAVDQECTSM----- 360
XX          |||          |||          |||          |||          |||
DB 513 dkpwley---pgynsgsdvsgmsypqsgllianwnltpydgilvrkkadsigmpidd 569
OY 361 -----DYVNSNTKLRPKGDMIGRIDALIDISTG-R 390
XX          |||          |||          |||          |||          |||
DB 570 pntkpgsggaegnandgfygylvtplwdgy-lymncnrrp-----ymaltadmkgqk 624
OY 391 TLMS-----VERAANYSPVLSTGGVLENG-CTDXYFRALSOETG 430
XX          |||          |||          |||          |||          |||
DB 625 vlwqhpjlgitarangpwjlpjglpweilgprnngsvlvqgllifgaatdnqiraidentg 684
OY 431 ETLWOTRLATVAGSQAISYEVDMQVVAIAGG 463
XX          |||          |||          |||          |||          |||
DB 685 kvwsavllpgsganpmtyeanghgyalmag 717

RESULT 10
ABG24430
ID ABG24430 standard; Protein: 443 AA.
XX
XX ABG24430;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #24421.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX

```

1

243 Q I E F R N W D Q E C T F E M M V T N --- V D V Q P S T E M E G L --- Q S I N P N A T G E R R V - L T G V P 263

Db 510 -lppvdd-----vtkygvdgqplsteeipfddkkrfennlkkpgeervkqgpe 558
 QY 264 -CRT-GTMMQFAEENGELMARADTNYOMLIESIDENGIVTNEADILKELDVEXVCTFE 321
 Db 559 gktlttpttknplqekvggeptekltkqvde--iteygggealkpghkdeifdpnqk 616
 QY 322 LGGHDWPS--AALNPDSS-ITFIPLNANCY-----DMAVNDGF--FTSMDEVNTSNVTK 370
 Db 617 gsgedvpgkqgvknppdtgtvvpvddvltkypvgdgplsteeipdkkrfennp---- 672
 QY 371 LPECKDMI 378
 Db 673 lkpqkerv 680

RESULT 12
 ABG25909
 ID ABG25909 standard; Protein; 948 AA.
 AC ABG25909;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #25900.
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR N-PSDB; AAS90096.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 56268; 103bp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human
 diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX Sequence 948 AA;
 SQ
 Query Match 4.8%; Score 122.5; DB 22; Length 948;
 Best Local Similarity 34.4%; Pred. No. 0.11; Indels 25; Gaps 5;
 Matches 33; Conservative 14; Mismatches
 QY 6 LMAA-----AGALAILAPAFRAQVTPVDELANPAGEMISYGOENYRHSPLTO 58
 Db 310 ltwaqndpgeingtltsadapaea-ispvddq-----dwpaygrnqegqrlspkq 360
 QY 59 ITTENYGOLOLWMA-----RGMPGRV--QVTPPL 85
 Db 361 lndnvnmlkeavfrtgvdkqndpgeitnevpl 396

RESULT 13
 ABG21573
 ID ABG21573 standard; Protein; 1510 AA.
 AC ABG21573;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #21564.
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR N-PSDB; AAS85760.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 51932; 103bp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 1510 AA;

Query Match
 Best Local Similarity 4.8%; Score 122.5; DB 22; Length 1510;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;

QY 6 LMAS-----AGALALAAFAFAQVTPYDELIANPAGEWISYGOENRHSPLNO 58
 Db 310 ltwagfndpgeingtltsadaipaea-ispvadq-----dwpaygrngqgrfslpkq 360
 QY 59 ITTENYGOQLYWA-----RGMPGRV--QVPL 85
 Db 361 lmadvnhlkeawfrtgdkvkgpndpgeitnevtpi 396

RESULT 14

AA83007
 ID AAG83007 standard; Protein; 910 AA.
 AC AAG83007;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3108.
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimerly WJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53857.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 PS
 XX Claim 18; Page 819; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II) given in AAG81454 to AAG83120 from Staphylococcus epidermidis.
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically listing of the present specification,
 CC sequences given in the sequence listing of the present specification,

CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 910 AA;

Query Match
 Best Local Similarity 4.8%; Score 122; DB 22; Length 910;
 Matches 92; Conservative 37; Mismatches 146; Indels 98; Gaps 21;

QY 26 TPYDELILANPAGEWISYGOENRHSPLTOIT--ENV-----GQLVWARGMQ- 76
 Db 231 tptt-----knpltekvgegeplekvkqvdeltetggeelkpgkdefdpnarkgsge 286
 QY 77 --PGR-----VOYTPLIHOGVMILANIVANGYVAGSTCOYSPGCFVSGHDSATG 125
 Db 287 dvpqkpvkknptdgtvtpvd-----vtkygp-----vdgdpitst 324
 QY 126 EEL--WENYFIP-----RAGEBDETWGNDYEARMWTG-----AMGQITYP 165
 Db 325 eelpfdkktreldnlpagtekvvgkqgepklitpttknpltekvgegeplekvkq 384
 QY 166 VTNLVHRTGAVGPA--SETQRTGRTLTGNTFFAAR--PDTEIYWRHQTLPNDMDQ 222
 Db 385 vdelvhygeelkpgkdefdpnarkgsgevdvqkpgknpdgtgevv-----ltpvdd--- 437
 QY 223 ECTFEMWVTV--VDVQSTMEGL-----OSINPNATG--ERRVLGVP--CKT--GTMQ 271
 Db 438 -----vtkyrvpvdgdpitsteeelpfdkktreldnlpagtekvvgkqgepklitptt 490
 QY 272 FDAETGEFLMADTYQNMITEIDENGIVTWEDALIKELDYDVCPTFLGGRWPS-- 329
 Db 491 knpltekvgegeplekvkqvdeltetggeelkpgkdefdpnarkgsgevdvqk 548
 QY 330 AALNPDGSIYFIP 342
 Db 549 gvknpdgtgevtvp 561

RESULT 15

AA82914
 ID AAG82914 standard; Protein; 696 AA.

AC AAG82914;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimerly WJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53764.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 PS
 XX

PS Claim 18; Page 763-764; 2188bp; English.

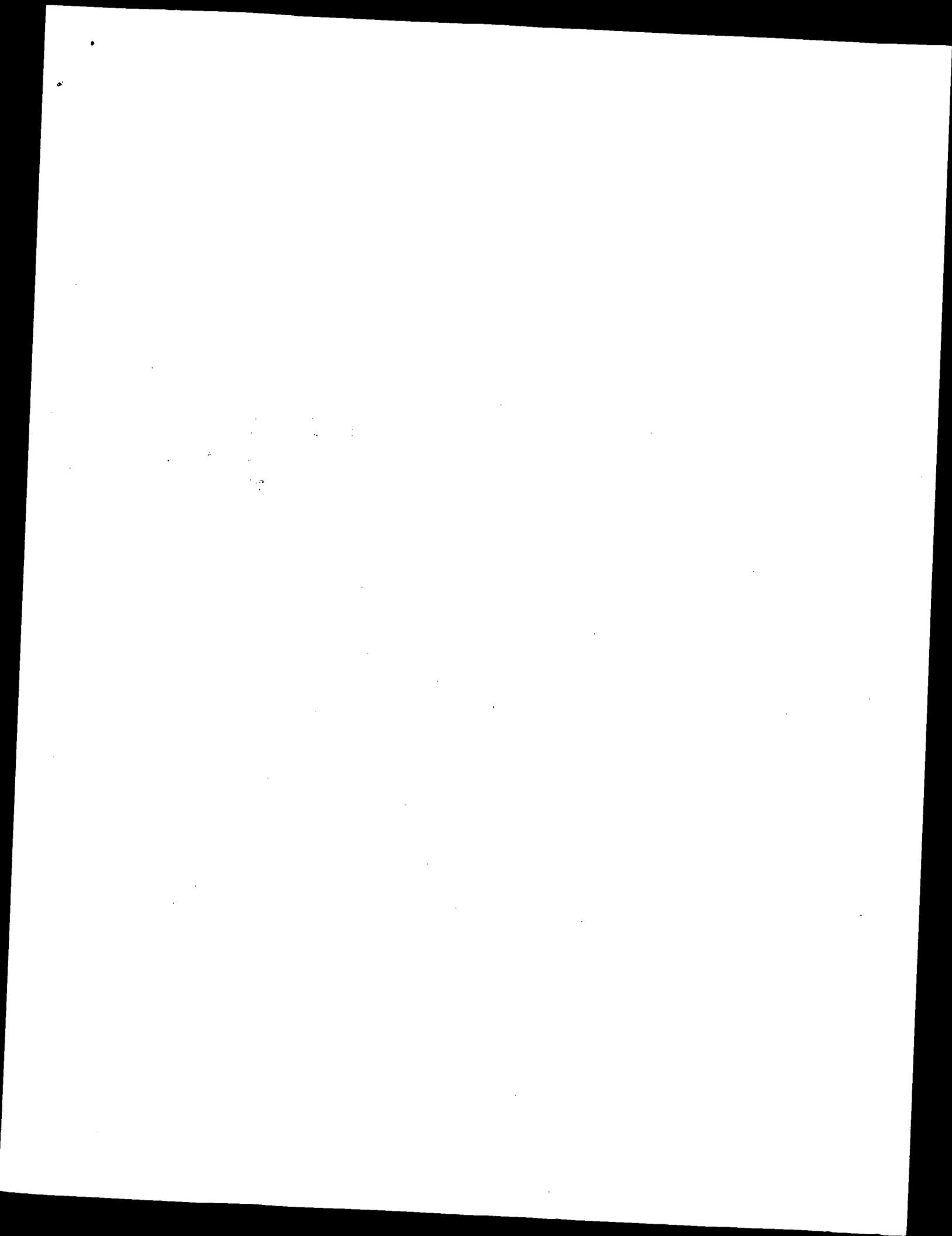
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AA68154 to AA683120, from staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are present for SEQ ID NO:4455 to 4464.
XX no sequences are present for SEQ ID NO:4455 to 4464.

SO Sequence 696 AA:

Query Match 4.8%; Score 121.5; DB 22; Length 696;
Best Local Similarity 23.5%; Pred. No. 0.087;
Matches 101; Conservative 48; Mismatches 147; Indels 133; Gaps 25;

OY 26 TPVTDLLANPPAGWISYQONQENYRHSPLTQIT---ENV-----GQLQLYWARGMO- 76
DB 103 tptt---knplgkvggeptektikpvdelleyggeikpghkdefpnapkysqe 158
OY 77 --PGK-----VOYTPLIHDGVMITANIYANGVIVASTCOYSPGCFVSGHDSATG 125
DB 159 dvpgkpgknpdgtgevtppvdd-----vtkygp---vdgdpilst 196
OY 126 EEL---KKNYFIP-----RAGEGDETWGNDYERKMTG-----AMGQITVDP 165
DB 197 eelpfdkkrelnpnlapgtkvkggeptektitptknplgkvggeptektikp 256
OY 166 VTNLVHVGSTAVGPA--SETQRTGPGTILGTNTREAVR--PDTEIWRHQTLPDNDQ 222
DB 257 vdeivhyggeelkpgkdefpnapkysgeadvpgkpgvknptdgev---lppvdd--- 309
OY 223 ECTFEEMWNTN---VDVOSTEMEGT-----OSINPNATG--ERRVLTGVP--CKT--CTMWQ 271
DB 310 -----vtkygpvdgdslsteeipfdkkrdefpnlapgtkvkggeptektitpt 362
OY 272 FDAETGEFLWARDNTYQNMIESIDENGIVTNEDEILKELDEYDVCPTFLGGRDWPSSA 331
DB 363 knplgtkvgegeptektikpvyde--lvhyggeqipgghkdef-----pnap 409
OY 332 L-----NPDSG-IYFIPLNNVCYDMAAVDOFTSMD-----VYNTSNVT 369
DB 410 vdsktevpgkpgvknptdgtvtpvddvtkygpvknplsteeipfdkkrvfnpd--- 466
OY 370 KLPGRKMI 378
DB 467 -lkpgeerv 474

Search completed: May 24, 2002, 10:19:16
Job time: 379 sec



Fri May 24 11:27:35 2002

walick-934-95.pep.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:14:47 : Search time 39.53 Seconds
(without alignments)
291.649 Million cell updates/sec

Title: WALICK-934-95.PEP
Perfect score: 2530
Sequence: 1 MKPTSLMASAGALALAAP.....GMOYVALAGGVSYGSLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues 231628

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/p/odata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/p/odata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/p/odata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/p/odata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/p/odata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/p/odata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413	16.3	754	4 US-09-296-284-4	Sequence 4, Appl1
2	407.5	16.1	738	1 US-07-985-458-3	Sequence 3, Appl1
3	404	16.0	720	4 US-09-296-284-25	Sequence 25, Appl1
4	279.5	11.0	739	3 US-09-136-251-2	Sequence 2, Appl1
5	105	4.2	1012	1 US-08-219-262B-10	Sequence 10, Appl1
6	105	4.2	1012	3 US-09-031-655-10	Sequence 6, Appl1
7	103.5	4.1	500	2 US-08-987-519-2	Sequence 1, Appl1
8	103.5	4.1	500	2 US-08-737-336-6	Sequence 1, Appl1
9	97.5	3.9	500	2 US-08-987-519-1	Sequence 2, Appl1
10	97.5	3.8	548	5 PCT-US93-10541-12	Sequence 19, Appl1
11	95	3.8	1012	1 US-08-216-276A-19	Sequence 2, Appl1
12	94.5	3.7	512	4 US-09-356-818A-2	Sequence 11, Appl1
13	94	3.7	1042	3 US-08-928-361B-5	Sequence 5, Appl1
14	94	3.7	1837	3 US-08-928-361B-5	Sequence 2, Appl1
15	94	3.7	5215	4 US-09-105-537-2	Sequence 2, Appl1
16	93.5	3.7	548	1 US-08-247-902A-2	Sequence 2, Appl1
17	93.5	3.7	583	4 US-09-401-476-2	Sequence 4, Appl1
18	93	3.7	380	3 US-08-971-782-4	Sequence 4, Appl1
19	93	3.7	380	3 US-08-971-782-4	Sequence 4, Appl1
20	93	3.7	459	4 US-08-971-782-2	Sequence 2, Appl1
21	93	3.7	459	4 US-09-309-026-2	Sequence 15, Appl1
22	92.5	3.7	1252	2 US-08-682-517-9	Sequence 9, Appl1
23	92.5	3.7	1252	2 US-08-682-517-9	Sequence 3, Appl1
24	92	3.6	606	2 US-08-883-534-3	Sequence 3, Appl1
25	92	3.6	606	2 US-09-204-764-3	Sequence 4, Appl1
26	92	3.6	812	1 US-08-446-794A-4	Sequence 4, Appl1
27	91.5	3.6	816	1 US-07-731-157A-4	Sequence 4, Appl1

28	91.5	3.6	816	1 US-08-229-444B-2	Sequence 2, Appl1
29	91.5	3.6	816	2 US-08-541-780-4	Sequence 4, Appl1
30	91.5	3.6	1612	1 US-08-169-927-2	Sequence 2, Appl1
31	91	3.6	551	2 US-08-793-229-32	Sequence 32, Appl1
32	91	3.6	551	3 US-08-285-957-32	Sequence 29, Appl1
33	90.5	3.6	322	1 US-08-216-276A-29	Sequence 2, Appl1
34	90.5	3.6	687	5 PCT-US91-09784-2	Sequence 5, Appl1
35	90.5	3.6	1721	3 US-08-700-651-5	Sequence 17, Appl1
36	90	3.6	484	1 US-08-476-276A-17	Sequence 3, Appl1
37	90	3.6	714	2 US-08-472-534-3	Sequence 12, Appl1
38	90	3.6	1012	1 US-08-219-262B-12	Sequence 12, Appl1
39	90	3.6	1012	3 US-09-031-655-12	Sequence 12, Appl1
40	90	3.6	1012	3 US-09-031-655-12	Sequence 2, Appl1
41	89.5	3.5	820	1 US-08-291-886-2	Sequence 2, Appl1
42	89.5	3.5	820	2 US-08-485-278-2	Sequence 3, Appl1
43	89.5	3.5	723	1 US-08-290-937B-3	Sequence 2, Appl1
44	89	3.5	723	1 US-08-290-937B-3	Sequence 2, Appl1
45	88	3.5	1012	1 US-07-944-943-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Lee, Eun-Hae
APPLICANT: Rhee, Sang-Ki
TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533, 0870000
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Glucobacter suboxydans
US-09-296-284-4

Query Match	16.38	Score 413	DB 4	Length 754
Best Local Similarity	23.88	Pred. No. 1.4e-31		
Matches 143	Conservative	72	Mismatches 222	Indels 164
				Gaps 15
OY	14 LALLAAPAFQVTPY-----	TDLLANPPAGEMISYGOENRHSPLQITTEVQ	66	
DB	17 LGGAAALPCATSPVALMEDGTATNADHPGDMWSYGRYSBORXSPLODITKDNASN	76		
OY	67 LQLVWARGMOPGKQVY-PTLHDSGVNYL-----	ANIVAN	99	
DB	77 LKLMHVDLDTNRKEGPTLVDSYMTATNWSKKALDANATGKLLMSYDPKVPGNADR	136		
OY	100 G-----		100	
DB	137 GCCDVTNNGAAYNMGKYFTFTDRLIALDAKTGLVSVTVPRKQALQHSRYTVDA	196		
OY	101 -----	VYAGSTCOVSPFCFVSGHDSATGEIMRNFTIPRAGEED	142	
DB	197 PRIAKGVITNGGAEEFGAR-FTYATDAETGKMDRFTVNPKNRKAASDVLMSK	255		
OY	143 --ETWGNDEARWMTG--	AMGQITVDPVNLVHYSSTAVGPASETORGTGGLTYGNT	197	
DB	256 APTWKGKMGKQGGGGTWDLSLITPDYDLVYLGVNSPNWYKFRSEGGNNLFLGS	315		
OY	198 RFAPRPTGELVWRHQLPDRNWDCEFTENAVYNDVQSTEMEGLOSINPAATGE-R	256		
DB	316 IVAINPTGKTVHHPETPDMDQMDYTSVQALADMPV-----	NGEMR	358	

QY 257 RVLTGVPCCKTGMQDAETGEFLMARDTYONMTESIDE-NGIYVNEAIIKEIDVEX 315
 Db 359 HVLVHAP-KNGFEIITDAKTKFKISGRPTYENMANGIDPVTGPNPNPALTUNGKPM 417
 QY 316 DVCPELGGDRMPSALINDPSGIYFIPLNVCY-----DMAVDEDFSMV 362
 Db 418 YGIGDGGHNFMAAMASPSOTKLYTIPAOQVPEYDDQKGFRAHDSNMLGLDMNKIGL 477
 QY 363 YNTSN---VTKLPGKDMIGRIDIDISTRTLSVERAANYSPVLTSGGGLFNGGT 418
 Db 478 LDNDQOHKADKAQFLKDLKGVIYAMDOKQAFFYDHHKGPWNGSLATAGVLFQGLA 537
 QY 419 DRYFALSOETGETLMQRLATVASGAISYEVDGMQYAI---AG-----GGVSYS 468
 Db 538 NGEFHAYDATGKDLFTTPAOSALIPVYTYANGQYAVAEVGGIYFPLGGLVARTS 597
 QY 469 G 469
 Db 598 G 598

RESULT 2

US-07-985-458-3
 Sequence 3, Application US/07985458
 Patent No. 5344777

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Fukaya, Masahiro;
 APPLICANT: Okumura, Hajime and
 APPLICANT: Kawamura, Yoshiya
 TITLE OF INVENTION: Structural Gene of Membrane-Bound
 TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
 TITLE OF INVENTION: Containing the Same And Transformed Acetic Acid
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fishaut, Holtz, Goodman & Woodward, P.C.
 STREET: 600 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
 COMPUTER: IBM PC compatible (NEC PC-9801 ES)
 OPERATING SYSTEM: MS DOS
 SOFTWARE: ASCII Form

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/985, 458
 FILING DATE: 19921203
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/658, 221
 FILING DATE: 20-FEB-1991
 APPLICATION NUMBER: 73440/1990

ATTORNEY/AGENT INFORMATION:

NAME: Goodman, Herbert
 REGISTRATION NUMBER: 17081
 REFERENCE/DOCKET NUMBER: 910134/HG

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)972-1400
 TELEFAX: (212)370-1622

TELEX: 236268

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 738 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FEATURE:

NAME/KEY: MATURE PEPTIDE
 LOCATION: 36 to 738
 IDENTIFICATION METHOD: N-terminal sequences of the
 IDENTIFICATION METHOD: purified protein having a molecular weight of about
 ORIGINAL SOURCE: 72,000
 ORGANISM: Acetobacter altoacetigenes
 STRAIN: MH-24
 PUBLICATION INFORMATION:
 AUTHORS: Tamaki, Toshimi;
 AUTHORS: Takemura, Hiroshi;
 AUTHORS: Fukaya, Masahiro;
 AUTHORS: Takemura, Hiroshi;
 AUTHORS: Takemura, Kenji;
 AUTHORS: Okumura, Hajime;
 AUTHORS: Kawamura, Yoshiya;
 AUTHORS: Mshiyama, Makoto;
 AUTHORS: Horinouchi, Sueharu and
 AUTHORS: Bepu, Teruhiko
 TITLE: Cloning and sequencing of the Gene Cluster
 TITLE: Encoding two subunits of Membrane-Bound
 TITLE: Alcohol Dehydrogenase from Acetobacter
 JOURNAL: Biochimica et Biophysica Acta.
 VOLUME: 1088
 PAGES: 292-300
 DATE: 1991
 US-07-985-458-3

Query Match

Best Local Similarity 16.18; Score 407.5; DB 1; Length 738;
 Matches 140; Conservative 68; Mismatches 219; Indels 163; Gaps 14;

QY 10 SAGALLAARAFQVTPVDE-----LLANPAEEMSTYCONENRHSPLQIT 61
 Db 17 TAGTICALLSGATVNASADGGATGATGALIHADHCHNMWYGRYSDDQSPLDQINR 76
 QY 62 ENVOQLOLVARQMOPGKQV--TPLIHGQVYL-----A 94
 Db 77 SNGNLKLAVIDLDITNRQGETPLVDGVMATVNMKAVADATGKLSYDPRPG 136
 QY 95 NI----- 96
 Db 137 NIDKSCDCTVNRGAANYNGVYGTGRLIALDAKTKLVASVNTTPPAELGKORSY 196
 QY 97 -----VANGYVAGSTCOYSPGCG--FVSGHDSATGELMRNFTIRAGECD----- 142
 Db 197 TYDAPRIAKGRVITGN--GGSEFGARGFVSADFETGKVDVREFVYVPPNRPDAADS 254
 QY 143 -----ETWGNDEYKRWMTGA-----WGQITDPVTLNHYGSTAVAPASETOR 185
 Db 255 VLMKAVQWTS-----PTGAMTQGGGGVWDSIYDPAVDLYLVGNGSPNNYKTR 307
 QY 186 GTPGGLYGTNTPFVAVRPTGEIVRHQTLERDWMDOECTFEEMVYVNVQVSTMEGLQ 245
 Db 308 SEGKGNLFLGSLVALKPEGTGEVWHFQETPMQWDETSPQINTIDLP----- 357
 QY 246 SINDNATERRVLTGVPCCKTGMQDAETGEFLMARDTYONMTESIDE-NGIYVNE 304
 Db 358 -----NETRHVIVHARKNGFFIIDAATGTFISGKNYVYVNMASGLDPKGRPIYNP 410
 QY 305 DALIKELDEYDVCPTFGGDRMPSALINDPSGIYFIPLNVCYDMAVDEDFSMV 362
 Db 411 DALYTLTGKEMYGIPDLGCHNFMAAMASPSOTKLYTIPAOQVPEYDDQKGFRAHDSNMLGLDMNKIGL 477
 QY 364 NTS---NVTKLPG-----KDMIGRIDIDISTRTLSVERAANYSPVLTSGGGLFNGGT 418
 Db 471 NLGIDMKVGIPIPSPEAKQAFKDLKGVIYAMDOKQAFFYDHHKGPWNGSLATAGVLFQGLA 537
 QY 413 LFGNGTDYFRALSOETGETLMQRLATVASGAISYEVDGMQYAI---AG-----GGVSYS 468
 Db 531 LFOGLANGEFHAYDATNGSDLFHFRAADSGIIPVYTYANGQYAVAEVGGIYFPLGGLVARTS 597

RESULT 3

US-09-296-284-25

Sequence 25, Application US/09296284A

Patent No. 6204040

GENERAL INFORMATION:

APPLICANT: Choi, Eun-Sung

APPLICANT: Rhee, Sang-Ki

APPLICANT: Lee, Eun-Hae

TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes

TITLE OF INVENTION: and Methods of Use Thereof

FILE REFERENCE: 1533.087000

CURRENT FILING DATE: 1999-04-22

NUMBER OF SEQ ID NOS: 87

SOFTWARE: Patent Ver. 2.0

SEQ ID NO: 25

LENGTH: 720

TYPE: PRT

ORGANISM: Glucobacter suboxydans

US-09-296-284-25

Query Match 16.0%; Score 404; DB 4; Length 720;
Best Local Similarity 23.5%; Pred. No. 9.8e-31;
Matches 137; Conservative 73; Mismatches 212; Indels 160; Gaps 15;

26 TPTVDELLANPPAGEMISYGNQENRHSPLDQTTENYGOLOLVARGMOPGVYV-TP 84
 5 TATTN---ADQHPGDMISYGRYSBORSPLDQTTKDNASNLKLAHMYLDNRGOEGTP 61
 85 LHDGMYL-----ANIVANG----- 100
 62 LIVDGVMTATTKMSKKKALDADATGKLMSYDPKVPNIADGCCDITVNRGAAYWNGKYTF 121
 101 -----VIVAGSTQYSP 112
 122 GTFDRLLADAKTKLVMSYVTVPREKQLOHQSIVYDGAIRAKKVIILNGAIEFGA 181
 113 FGCYVSGHDSATGELMRNYFIRAGEEGD-----ETNGNDYEARMMNG--- 156
 182 RG-EVTAIDAEFGKMDREFYVNPNDKPDGAASDVLSKAYPLPWGSGAMKQGGGGT 240
 157 AMGOITDPTNLVHGSTAVGASFGTGTGNTFRFAVRPDTEIWRHQTLP 216
 241 VMSLITPYVDIYLLGVGSGSPWNTKFRSEKGNLPLGSLVAINPDTGKYVNHFOETP 300
 217 RNMPOQECTEFEMAVTNDVQPTSEMGLOSINPNAATGS-RVYLTGVCKTGTMQFDAE 275
 301 MDQMYTISVQIIMALDMPV-----NGEMRHVLAHP-KNGFTYIIDAK 342
 276 TGEFLMARDNTYONKIESIDE-NCIYTVNEDAILKELDEVYDVCPTFLGGRMPSSALNP 334
 343 TGFYISGRKYTYEMNANGLDPTVGRPNYNDPALMTLNGKRPWGIIPDGLGHNFAAMAYSP 402
 335 DSGIYPIPLANNCY-----DMAVDOETSMYVITSN---VTKLPPGDM 377
 403 QTVLITPAQOVPFYVDQKGFKAHDSNMLGDMNKIOLDNDPQKADKAOFLDL 462
 378 IGRIDAIDISTGRITMSYERAAANSPVLSGSGVLTNGGTDYRFALISOETGETLMQTR 437
 463 KGMIVAMDPOKQAAFTVDHKGPMNGGLLAVAGVLEOGLANGEFNAHYATTKDLEFTRP 522
 438 LATYASGOAISYVDGMOYVAI---AG-----GVSYSGS 469
 523 AQSATIAAPYTYTANGLQYVAVGVGMIYFPLGVAFRSG 564

RESULT 4

US-09-136-251-2

Sequence 2, Application US/09136251A

Patent No. 6127156

GENERAL INFORMATION:

APPLICANT: HOSHINO, Tatsuo

APPLICANT: MIYAZAKI, Taro

APPLICANT: OJIMA, Setsuko

APPLICANT: SHINJOH, Masako

APPLICANT: TOMIYAMA, No. 61271561bunmi

TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE

FILE REFERENCE: D-Sorbitol Dehydrogenase Gene

CURRENT FILING DATE: 1998-08-19

EARLIER APPLICATION NUMBER: EP 97114432.4

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patent Ver. 2.1

SEQ ID NO: 739

LENGTH: 739

TYPE: PRT

ORGANISM: Glucobacter suboxydans

FEATURE:

NAME/KEY: SIGNAL

LOCATION: (1)..(24)

US-09-136-251-2

Query Match 11.0%; Score 279.5; DB 3; Length 739;
Best Local Similarity 20.8%; Pred. No. 1.6e-18;
Matches 144; Conservative 70; Mismatches 165; Indels 313; Gaps 33;

23 AQVPEVDELLANPPAGEMISYGNQENRHSPLDQTTENYGOLOLVARGM-OPG 78
 86 SQVPAMAPQGSANPARDPDVAYGRDHQRTYSPLSITPENSKLVAFFVHYSRPPG 145
 79 KY-----QVTP-L-IHDGMYLANIVANGVIVAGSTQYSPFCFPGSGHDSATGEELMR- 130
 146 QVNMKAETPLIKYVDS-LYTCSAM-NDIILK-----DPAIKQIMWRN 187
 131 -----NYF-----IPRA----- 137
 188 VDKYHSIPTAACKGVYTFSSVPSGQCHNRLIGETLMDRLIADAEFGDPCPNFGH 247
 138 -----GE-----EDDETW-----GNDYEARMM 154
 248 GGOVNLMOGLIGESVGFVSMAPRPVINGVVVNHVYLDQRRWADSGVIRGDAESGKF 307
 155 TGAM-----GOITDPTNLVHGSTAVGASFGTGTGNTFRFAVRPDTEIWRHQTLP 198
 308 VMADVDVNSGRSOPAY-RVYITVAVERRIPGLPDRORG-----GSRLLRPPRNSADY 359
 199 -FAVRPD-----TGEIWRHQTLPDNDNDQCTEFEMAVTNDVQPTSEME 242
 360 YSALRSDAENKYSAAVYALDVKTGSPRWVFTAHNDVDYDYGSOATL-----MDMP 411
 243 GLQSLNPNAATGERRVLTGVPCKTGTMQFDAETGFEFLMARDNTYONKIESIDEN----- 297
 412 G-----PDGQTVPALIM---PTRKGQTFVLDRTGK-----PLPVEERPAAP 452
 298 GIV-----TVNEDAI-----LKEID----- 312
 453 GVTPGDPSPQTPMSVGMALRYVDLKETDMGMSPTDOLFRCIRFRANVYGETTPSV 512
 313 -----VEYDVCTFLGGRDMPSSALNPDSGIYFIRLN-ANCYDMAVDOETSM----- 360
 513 DKPWLEY---PQYNGSDMGMSYDPOSGLILANNITTMQYDQVLTBRKADSLGIMPID 569
 361 -----DVYNTSNVTKLPCKDMKIGRIDAIDISTG-R 390
 570 PNFKPGGGAEGNGAMDGPYGIYVTFPMQY-TGMMCNRP---YGMITLIDKHKQK 624
 391 TMS-----VERAANSPVLSGSGVLEFG-CTDRYFRALISOETG 430
 625 VLMQHPILGTARANPGLPTGLPWEITPNNGSSVYGGGLIFGATDNOIRAIIDEHTG 684
 431 ETLMQTRLATVASGOAISYVDGMOYVAIAG 462


```

Db 61 PGFP--GSVGAHTYLOSNGSYQFDO-----MLTLAONTLPVSYNVCCLVSRSLTVSSLT 113
QY 241 MEGLOSINP--NATGERRVLGVCKTGTMMQFPAEGETFIMARDTNYONMTESTI----- 294
Db 114 PGVYALNGTINAVT-----FQGSISEL--TDYSYGLMSATANIN 152
QY 295 DENGIVYVNEADALKEDEYEDVCPFLGGRDMPSAALNPDGSIYFPLNNVC----- 347
Db 153 DRIGVAVLGEVYVLSLPTSYDLSYVRLGD-PIPAAGLDP-----KLMATCSDRPR 204
QY 348 -YDMAVDO-EF-----TSMOVYNTSNVTKLPKDKMI-GRIDALDISGRILM- 393
Db 205 YTVTAADDEYOFSSQLIPSGVKITLFTANIDALISLVSGLIFSGVTHSLEVDVITIF 264
QY 394 -----SVERAANYSPLVSTGGVLEN--GQTDRIYFPAALSGEFTLMQTRLATYAS 443
Db 265 ICFDGEVYVKAIVUDEFLTGTNNLVPFNLGPTSETTOPITSMKLEVVYTKRGGT--A 322
QY 444 GQALSTEVDGMQYVALAG 462
Db 323 GDPISWVSGTLAVTVGG 341

RESULT 7
US-08-987-519-2
; Sequence 2, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Ludmeyer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987,519
; CURRENT FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 500
; TYPE: PRP
; ORGANISM: Human Papillomavirus
; US-08-987-519-2

```

```

Query Match 4.1%; Score 103.5; DB 2; Length 500;
Best Local Similarity 21.4%; Pred. No. 0.14; Indels 79; Gaps 11;
Matches 52; Conservative 30; Mismatches 82
QY 36 PPAGEWISYGCNDENYRHSPLQITTEVNGQLVWARGMOPGKVOTPLIHD----- 88
Db 158 PPLGHEHMGCKGKCTN--TP-----VOAGDCPPLRLITSVLQDGDVDTG 200
QY 89 -GVYLIANIVAN--GVIVAGSTCOY-----SPFG--CFVSGHDSANTGEELMRN 131
Db 201 PGAMNFALQNKSDVPIDIGTCKPYDLOMADPYGDRLEFFL-----REOMFAR 254
QY 132 YFIPRAGEGDETWGNDYFARMWGANGQITTYDPTNLVHGSYAVGPASETQRTPGGT 191
Db 255 HFNRRAGEVEEV-----PDLILIKSGNRTSVASSIYVTPSGS 294
QY 192 LVGTNTRFAVRPDTGEIYWRHQTLPRLN--WDOECTFEMAVTVNDVQSTENEGLOSIN 248
Db 295 LVSEAOLEFNKP-----YWLQKAGCHNNGICWGN--QLFVTVVDTTSTNNLTICASVT 345
QY 249 PNA 251
Db 346 TSS 348

RESULT 8
US-08-737-336-6
; Sequence 6, Application US/08737336
; Patent No. 6013262
; GENERAL INFORMATION:

```

```

APPLICANT: FRAZER, Ian
APPLICANT: ZHOU, Jian
TITLE OF INVENTION: RECOMBINANT PAPILLOMA VIRUS LI
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,336
FILING DATE: 16-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU95/00292
FILING DATE: 17-MAY-1995
APPLICATION NUMBER: AU PM 5667
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 65064/118
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-336-6

```

```

Query Match 4.0%; Score 100; DB 3; Length 532;
Best Local Similarity 21.6%; Pred. No. 0.33; Indels 122; Gaps 18;
Matches 68; Conservative 37; Mismatches 88
QY 3 PTS--LLMASAGALALAAPAFQVTPYTDLELANPPAGEWIS-----YGCNDENYH 53
Db 122 PTQRLVWACTGL-----EYGRGQPLGVGSGHPLNKYDVENSGSG 164
QY 54 SPLTQITTEWVG-----OLOLV-----WARGMO-----PGKVOTPLIH 87
Db 165 GNPQDNRVNGADYKOTQLCVNGCAPLGEHMGKCKCTNTPVOAGDCPPLRLITSVLQ 224
QY 88 D-----GVYLIANIVAN--GVIVAGSTCOY-----SPFG--CFVSGHDSA 123
Db 225 DDDWVDTEFGAMNADLOTKNSDVPIDIGTCKPYDLOMADPYGDRLEFFL----- 278
QY 124 TGEELMRVYFIPRAGEGDETWGNDYFARMWGANGQITTYDPTNLVHGSYAVGPASET 179
Db 279 KREOMFAHFENRAGEV-----EVPDLILIKSGNRTSVAS 315
QY 180 ASETQRTPGGTLLYTNTNRFVRPDTGEIYWRHQTLPRLN--WDOECTFEMAVTVNDVQ 236
Db 316 -SSIYVTPSGSLVSEAOLEFNKP-----YWLQKAGCHNNGICWGN--QLFVTVVDT 365
QY 237 PSTEMEGLSINPNA 251
Db 366 RSTNNLTICASVTSS 380

RESULT 9

```

```

US-08-987-519-1
; Sequence 1, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Ludmerer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; CURRENT APPLICATION NUMBER: US/08/987,519
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human Papillomavirus
US-08-987-519-1

```

```

Query Match
Best Local Similarity 3.9%; Score 97.5; DB 2; Length 500;
Matches 66; Conservative 35; Mismatches 99; Indels 109; Gaps 16;

```

```

QY 3 PTS--LWASAGALALALAPAFQVTPVDELANPAGEWIS-----YQ--QNGENY 51
DB 90 PTTORLWACTGL-----EVGRQPLGVGVGHPPLNKTDVENSNGY 132
QY 52 RHPSLQITTEWNG-----QIQLV-----WARGMO-----PGKVQVTP 85
DB 133 GGN--GDNRYNVGMDYKOTOLCMVGCAPLGEHMGKTKTOSNTSVNGCCPLELITSV 125
QY 86 IHD-----GYMLANIVAN-----GYVAGSTCOY-----SPFGCVFSGHDSANG 191
DB 192 IOGDWVDTGFGAMNADLQTKSDVPLDGTGCKTPDYLOMAADPYGDRLEFY--LTK 249
QY 126 EELMRVFTIPRAGEGSDDEWNGNDYEARMTGAMGQIYDVTNLYHGSTAVGPAEFTOR 185
DB 250 EQMFAHFENRAGTGEV-----PDDLTVGKNNKSSVASSITV 289
QY 186 GTPGGILGTNTREAVRDPDTEIWRHQTLPDN---WQECFEMATNTNDVOPSTEME 242
DB 290 HTPSSGLVSSSAOLFENK---YWLQKAGHNGICWGN---HLFVTVDTRSTMT 340
QY 243 GLQSLNPA 251
DB 341 LCASVSKSA 349

```

```

RESULT 10
PCT-US93-10541-2

```

```

; Sequence 2, Application PC/TUS9310541
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: TRANSGLOUTAMINASE GENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: USA
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10541
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126119
; FILING DATE: 23-SEP-1993
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 07/969702
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Eugenia S. Hansen
; REGISTRATION NUMBER: 31966
; REFERENCE/DOCKET NUMBER: 33590CIRCPC, OMRF B-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-10541-2

```

```

Query Match
Best Local Similarity 3.8%; Score 95.5; DB 5; Length 548;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

```

```

QY 118 SGHDSATGEELMWRNPIPRAGEGDETW-----GNDYEARMTGAMGQITPDPTNLVH 171
DB 17 NGRDHHTA-DLGRKELVRRQ---PFWLTHFGNRYEA-----SVDSLTFSTVY 63
QY 172 YGSTAVGPAEFTORGTGGTILGNTNPRVAPDTEIWRHQTLPDNNDQCFEEMVY 231
DB 64 -----GPAHQEA-----GKARFPLRDAVEGDMTATVVDQ---QDCITSLDT 105
QY 232 NVDVOPSTEMEBLOSINENATGER-----VLTVGPCKITGM--QDAETGEF-L 280
DB 106 T-----PANAPGLYRLSLERASTGYOGSSVFLGHILLFNAMCPADAVYLDSEEROEYVL 161
QY 281 WARDTNYQ---NMIESIDENGIVTNEDALIKEDVEYDVCFFL--GGRDMPSEA----- 331
DB 162 TQGGFTYOGSAKFIKNIPW--FGQFEDGIDLICLLIDVMPKFLKNAGRCSRSPY 219
QY 332 -----LNDSGIYFPLNNCYDMAVDOFTSMOV---YNTSNTKLPCKDKMI 378
DB 220 VGRVSGMNVNCNDQGVLLGRMDNNYGDVSPMSWIASVDILRRKKHGCORVKYGCWV 279
QY 379 GRIDAIDISTGRTLMSVERAANYSPVLSTGGVLFNGGTDRTFR---ALSOETGETLM 434
DB 280 --PAAVACTVRLGIGIPRTVYTNVNSAHDNSNLT-----EYFRNEFGEIOGDKSEMIW 332

```

```

RESULT 11
US-08-216-276A-19

```

```

; Sequence 19, Application US/08216276A
; Patent No. 5595912
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID
; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
; TITLE OF INVENTION: ASSOCIATED WITH US IBV VARIANTS, VECTOR CARRYING DNA
; SEQUENCES, HOST CARRYING CLONED VECTOR, DEDICED AMINO ACID
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```


Fri May 24 11:27:35 2002

walick-934-95.pep.rai

APPLICATION NUMBER: US/08/216,276A
FILING DATE: 23-MAR-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/083,784
FILING DATE: 28-JUN-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/519,202
FILING DATE: 04-MAY-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/227,311
FILING DATE: 02-AUG-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kelder, Steven B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-276A-19

Query Match 3.8%; Score 95; DB 1; Length 1012;
Best Local Similarity 23.8%; Pred. No. 2.9; 132; Indels 88; Gaps 20;
Matches 82; Conservative 43; Mismatches 132; Indels 88; Gaps 20;

178 GPAS-----ETGRTGCTLYGNTFRFAVRPDGTGEIYWRHQTLPD--N 219
Db 24 GPASIPDDLEKHLRSTSTYNTVGDGSLVFFGEP--GSLVGHANTLQSNQNYK 81
QY 220 WDECTFEEMAVTNDVDPSTMEGL--QSNPNAATGERRVLGVPCKTGMQPDDET- 276
Db 82 FDQ-----MLTFAOMLPSPSYNCRVRSRLVRSST--LPGGVYALNGT--NAVTF 129
QY 277 -GELIARPDYONNIESI--DENGIVYNEALIKELDEYDYVDCPTFLGGRDWPSSA 331
Db 130 QGSISETLTVSYNGMSATANNDKIGNYLVGEGVTVLSLPTSYDGLYRLDD-PIPAIG 188
QY 332 LNP-----DSG---IYFIPLNVCYDMAVDOFTSMQVYNTSNV-----T 369
Db 169 LDPKVVAVTCDSDRPVYTI-----TAADYQFSSQ--YQYGVTTITLFSANIDAIT 238
QY 370 KLPFGKDMIGRIDIDISTGRTLSV--ERAAVNSPVLTGGVLFNGGTDRI--FRAL 425
Db 239 SLISVGEIVFTSVHSLVLTATITLIGFGSNAVITRAVAANG--LTTGDNLMPEMLV 295
QY 426 --SQETGETIMQTRLATVAS-----GQALSYEVDCMQYVALAGG 462
Db 296 IPINETOPTISIKILEIYTSKSGGQEGDDQMSGSLAVTTHGG 340

RESULT 12
US-09-356-818A-2
Sequence 2, Application US/09356818A
Patent No. 6190896
GENERAL INFORMATION:
APPLICANT: Fraaij, Bassem M.
TITLE OF INVENTION: Active Human Cellular Transglutaminase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bassem M. Fraaij
STREET: 2401 N. Stat. Dr.
CITY: Stillwater
STATE: Oklahoma
COUNTRY: United States of America
ZIP: 74075

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5, 1.44 Mb High Density Diskette
COMPUTER: IBM AT Compatible
OPERATING SYSTEM: Windows NT
SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/356,818A
FILING DATE: July 19, 1999
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 2451,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (405)652-7252
TELEFAX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 512
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-356-818A-2

Query Match 3.7%; Score 94.5; DB 4; Length 512;
Best Local Similarity 21.4%; Pred. No. 1.1; 148; Indels 87; Gaps 19;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

QY 118 SGHDSATGEELMRNYFIPRAGEDETW-----GNDYEAHWMGTGANGQITTPYTNLVH 171
Db 17 NORDHHTA-DLCREKLVRRGQ--PFWLTIFEGRNEA-----SDSLTFSVYV-- 63
QY 172 YGSAVGAASEFGRTGTLGNTFRFAVRPDGTGEIYWRHQTLPDNDDECTFEEMAVT 231
Db 64 -----GRAPSOEA-----GTRARFPLRDAVEEDMTFATVVO--QDCTLSOLT 105
QY 232 NVDVDPSTMEGLQSNPNAATGER-----VLTVGPKGTGM--QDAETGEF-L 280
Db 106 T--FANAPIGLYRSLBASIGQSSFVLGFIILFNAMCPADAVYDSEERQEVVL 161
QY 281 WARDTNYQ--NMIESIDENGIVYNEALIKELDEYDYVDCPTFL--GGRDWPSSA-- 331
Db 162 TQGFYTGSAKFKIKNIPNN--FGQFEDGILDICLLIDVNPFLKRNAGRDCSRSSPY 219
QY 332 -----LNDPSGIVFIPLNVCYDMAVDOFTSMQVYNTSNVTKLPFGKDM 378
Db 220 VGRVVGAVNCDQSVLLGRMDNNYGDGVSFMSWIGSDILRRKRNHGGQRYKGCQCV 279
QY 379 GRIDAIDISTGRTLSVYERAAVNSPVLTGGVLFNGGTDRI--ALSOETGETIM 434
Db 280 --FAAVACTVLRCLGTPTRVVTNINSAHQNSNLLI-----EYFRNEFGELGDKSMIV 332

RESULT 13
US-08-928-361B-11
Sequence 11, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIRSA
STREET: 365 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: VERNY, Hana
 REGISTRATION NUMBER: 30,518
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1042 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-11

Query Match 3.7%; Score 94; DB 3; Length 1042;
 Best Local Similarity 20.2%; Pred. No. 3.9;
 Matches 131; Conservative 66; Mismatches 212; Indels 240; Gaps 33;

25 VTPVDELNPAGWISYQNOE-----NYRHSPLQITTEVNGOL----- 67
 Db 301 IDPTTG-LPENNPTGLHINPTNNMTDSSFAKAYKAYNSGKTDVYGLPVDITGLPK 359
 68 -----OLV-----WARGMOPKGVVTP-----LIHDEV 90
 Db 360 DVSDDIPFNSTGELVDSTGKPINNYTAGIVSGKRGLPIDENGNLFPSTKLPIDGN 419
 QY 91 MYLANIVANGVIVASTQYSPF-GCFVSGHDSATGEELMRNYFIRAGEGDETWGNDY 149
 Db 420 NQLVNPTNSTVSGSTSTKPKGIPVNGGVVDE-----AKQADK--GKD- 467
 QY 150 EARMWTGAMGQITDPTNLYVHGSTA--VGPASETORGTPG----- 189
 Db 468 --GLIVPTNINNDPVTNLYVHGSTA--VGPASETORGTPG----- 189
 QY 190 ---GLVGTNTRFAVRPTGIVRHQTLPRN--WDOCTFEEMVNTN-----VDVQ 236
 Db 524 EITGKPVDTVGLPYDSTGELIDPATKLPFGSVAGDELLEVTINITTDEVTGLPDLR 583
 QY 237 ---PTEMEGIQS-----INP-----NATG-----ERRVLTGVPCKTGM 269
 Db 584 TGLPRDPSVGLPOLNCTGLVDPSNKKRPIGSHSGFINGTSEOSHEKDPSTGLPDPNG 643
 QY 270 WQFD-----AETGEFLMARDNTQNM--ESIDENGIVTVEADILKEIDV-----E 314
 Db 644 LPFEDSGSLINPETGDKIQSHSTPMVPGKQGGENGIMPEQ--ILEALNKLPTSGNE 702
 QY 315 YDVCP-----TFL--GGRDMPASAL----- 332
 Db 703 VNISPSSDADVPDRPTMTNMKISGQTFQVQDGGKTIIPGSAASVHTALGPTQTDPTTG 762
 QY 333 ---NPDSGIYITPLNNWCYDMMAYDQ-----EFTSMQVNTSNVTKLPKGDMLGRIDAI 384
 Db 763 LPSDSTGILPFGFNVNLPDQNGEOLKSGSVPLSVLYKEKNITVEAYGL-----PV 815
 QY 385 DISTGRTMSVERAANYSPVLTG-----GGVLFNGTDRYRA-----LSETGE 431
 Db 816 DKTGTFIDPI--SILPFANGLIDIPISGKYFSSINGFISGAGSOSKSSDSEGN 870
 QY 432 TLMGT-----RLATVASGQAISEYVGMQVVALAG-----GGV 464

Db 871 PIDEPTNMPDKTKGLIDPESGIALDINSVSV-FATVPGTAAPKKGCV 918

RESULT 14
 US-08-928-361B-5
 ; Sequence 5, Application US/08928361B
 ; Patent No. 6071518
 ; GENERAL INFORMATION:
 ; APPLICANT: Petersen, Carolyn
 ; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 ; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 ; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PETERS, VERNY, JONES & BIRSA
 ; STREET: 385 Sherman Avenue, Suite 6
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306-1840
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/928,361B
 ; FILING DATE: 12-SEP-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/026,062
 ; FILING DATE: 13-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: VERNY, Hana
 ; REGISTRATION NUMBER: 30,518
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-324-1677
 ; TELEFAX: 650-324-1678
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1837 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-928-361B-5

Query Match 3.7%; Score 94; DB 3; Length 1837;
 Best Local Similarity 20.2%; Pred. No. 9.7;
 Matches 131; Conservative 66; Mismatches 212; Indels 240; Gaps 33;

25 VTPVDELNPAGWISYQNOE-----NYRHSPLQITTEVNGOL----- 67
 Db 1096 IDPTTG-LPENNPTGLHINPTNNMTDSSFAKAYKAYNSGKTDVYGLPVDITGLPK 1154
 68 -----OLV-----WARGMOPKGVVTP-----LIHDEV 90
 Db 1155 DVSDDIPFNSTGELVDSTGKPINNYTAGIVSGKRGLPIDENGNLFPSTKLPIDGN 1214
 QY 91 MYLANIVANGVIVASTQYSPF-GCFVSGHDSATGEELMRNYFIRAGEGDETWGNDY 149
 Db 1215 NQLVNPTNSTVSGSTSTKPKGIPVNGGVVDE-----AKQADK--GKD- 1262
 QY 150 EARMWTGAMGQITDPTNLYVHGSTA--VGPASETORGTPG----- 189
 Db 1263 --GLIVPTNINNDPVTNLYVHGSTA--VGPASETORGTPG----- 189
 QY 190 ---GLVGTNTRFAVRPTGIVRHQTLPRN--WDOCTFEEMVNTN-----VDVQ 236

```

Db 1319 EITGKRVDTVITGLPYDPSTGELIDPATKLPISGVADELITVLTDEVTGLPIDE 1378
OY 237 ---PSTEMGLOS-----INP-----NAATG---ERRVLTGYPCCTGM 269
Db 1379 TGLPRDPSGLPOLNGTIVDPSSNKKPIPGSHSGFINGTSGSESHEDPSTGKPLDPNG 1438
OY 270 WQPD-----AETGEFLMADNTYQNM--ESIDENGJYTVEDAILKELDV-----E 314
Db 1439 LPDEDESGSLINPETHDKIQSGSHSGTFMFPVPGKPGEGNGIMTEQ--ILEALNLPISNE 1497
OY 315 YDVC-----TFL--GGRDWPSAL-----332
Db 1498 VNISPRESSDAVDRPTNTWNNKISGCTQVNDGKKTIGSAASYIHTALGTPOTDPTG 1557
OY 333 ---NPDGIFIPLNNVCYDMAADQ---EFTSMDYNTSNVTKLPKPGKMDIGRIDAI 384
Db 1558 LPSPSTGLFPIGFVNLVDPQGEQIKGSVPVSLYXKNITVEAAYGL-----PV 1610
OY 385 DISTGRILMSYERAANYSPVLTG-----GGVLFNGGTDTRYFRA-----ISOETGE 431
Db 1611 DPKTGFPIDPI---SYLPPAKNGELIDPISGKTFSGSIAGPISGKAGSOSKSDSEGN 1665
OY 432 TLMOT-----RLATVASGOAISYEVDGMQYVALAG-----GGV 464
Db 1666 PIDPSTNMPYDPKTKLIDPESGIAIDNSYGV-FATVPETAAPKKGV 1713

```

```

RESULT 15
US-09-105-537-2
: Sequence 2, Application US/09105537A
: Patent No. 6285202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D. H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600,438051
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 5215
: TYPE: PRT
: ORGANISM: Streptomyces venezuelae
US-09-105-537-2

```

```

Query Match 3.7%; Score 94; DB 4; Length 5215;
Best Local Similarity 21.3%; Pred. No. 53;
Matches 114; Conservative 49; Mismatches 160; Indels 212; Gaps 29;

```

```

OY 67 LQLVWARGMOPGKYQVTPPLIHGVMYLANIVANGVYAGSTCOYSPFCFVSGHDSATGE 126
Db 1992 LELALAGDHGAVRVELELLEAPLVPERGAARIQVG-----VSGDGESPAQ 2039
OY 127 ELMRNYFTPRAGEEDET---WGNDYEARMTGAMQITVDPVTLVH-----YGSTAVG 178
Db 2040 RIFGVYSTPDSGDTDDAPREM---TRHYSVILGE--GDPATESDHPGTDGSAAMP 2092
OY 179 PASETORGTPG-----GTLGNTNRFVAPRPTGELIWRHOTLPDNDMDCECTFEMMY 230
Db 2093 PAAATATPLDGYDRIALGTYGP---AFQGLTG--LW-----RDGAD-----2131
OY 231 TANDVO--PSTEME--GLOSINP-----NAATGERRVLTGVPCKTGTMMQPD 273
Db 2132 TLAETIRLPAQHSAGLFEVHPALDAILHPVILEGNSAAG-----AC-----D 2175
OY 274 AETG-----EFLMADNTYQNM-----MIES--293
Db 2176 ADTDTATRIPLPAMAGVTLHAEGATALKVRITPTGPTVTLRLDTTGAVATVESLTL 2235

```

```

OY 294 -----IDE-----NGIVTVNEDAILKELD--VEYDVCPT 320
Db 2236 RAVAKDRLTGTAGRVDALEFTVWTFETGPBPAGGAVEVEELVDLAGLDLVLGADV 2295
OY 321 FLAGRDW-----PSALNPDGSIYFIPLANNVCYDMAVDOEFTSMDYNTSNVTKLP---372
Db 2296 VLRADRWTLIDGDESSA-----ARTAVRRLAIYOEFLSPRRDGSRLVCYTRGA 2344
OY 373 ---PGKMDIGRIDAIDISTGRILMSYERAANYSP---VLTGGGYLFGNGCTDRYFRA 424
Db 2345 VAALPGEV-----TSLATG-PLWGLVRSAGSENGRFLFLIDIGEGERDGADELIRA 2397
OY 425 -----LSOETGETLMQTRLATVA-----SGQAISEYVDGMQYVALAGGV 464
Db 2398 ATAGDEQLAARDGRLL-APRIARTALSSSEDTAGADRFGPDGIVLYTGGTGL 2451

```

```

Search completed: May 24, 2002, 10:20:03
Job time: 316 sec

```

Fri May 24 11:27:35 2002

walick-934-95.pep.ra1

OM protein - protein search
Run on: May 24, 2002, 10:15:22 ; Search time 54.76 seconds
(without alignments)
828.235 million cell updates/sec

(without alignment) updates/sec
828.235 Million cell

Title: WALICA_059_2
Perfect score: 2530
Sequence: 1 MKPTSLIMASAGALLAAP.....GMQYVAIGGVSYSGINS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 9006666

Total number of hits satisfying

Minimum	DB seq	length:	20000000000
Maximum	DB seq	length:	20000000000

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Listing first 45 summaries

Database :

```

1: pir2:*
2: pir3:*
3: pir4:*
4:

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	427	16.9	601	2	E95863	alcohol dehydrogen
2	422.5	16.7	708	2	S52317	guinohemoprotein e
3	421	16.6	742	2	RA9340	alcohol dehydrogen
4	409	16.2	742	2	S90326	alcohol dehydrogen
5	407.5	16.1	738	2	S14270	alcohol dehydrogen
6	386.5	15.7	626	2	U00706	methanol dehydroge
7	359.5	14.2	573	2	S68591	guinoprotein alcoh
8	357	14.1	623	2	B83399	glucose dehydrogen
9	299.5	11.8	796	1	JV01077	glucose dehydrogen
10	299.5	11.8	796	2	H90604	glucose dehydrogen
11	299.5	11.8	796	2	H85495	glucose dehydrogen
12	295.5	11.7	796	2	AG0523	glucose dehydrogen
13	289	11.4	801	1	S00943	glucose dehydrogen
14	251.5	9.9	808	2	OPKEX	glucose dehydrogen
15	237	9.4	803	2	F83360	probable guinate d
16	223	8.8	809	2	B98314	hypothetical prote
17	223	8.8	809	2	A12968	glucose dehydrogen
18	216	8.5	778	2	G98221	glucose dehydrogen
19	216	8.5	778	2	A13064	hypothetical prote
20	209.5	8.3	809	2	A41378	guinate-shikimate
21	209.5	8.3	809	2	A55547	polvinyl- alcohol
22	187	7.4	639	2	JC4881	polyvinylalcohol d
23	144.5	5.7	534	2	BA82580	probably secreted
24	133	5.3	839	2	D97013	probably serine/th
25	127.5	5.0	5627	2	C75264	hypothetical prote
26	124.5	4.9	5627	2	C83339	probable hemolysin
27	119.5	4.7	2535	2	AC0304	serine/threonine p
28	119	4.7	407	2	H69064	hypothetical prote
29	119	4.7	424	2	T29127	

30	119	4.7	927	2	F64512
31	118.5	4.7	386	2	A82284
32	118.5	4.7	799	2	T78989
33	117.5	4.6	1197	2	D82686
34	115.5	4.6	5925	2	A93953
35	115.5	4.6	1568	2	A66036
36	115.5	4.6	1588	2	H91188
37	115	4.5	1443	2	S23771
38	115	4.5	1044	2	F93375
39	114	4.5	796	2	T33962
40	114	4.5	1441	2	B86807
41	112.5	4.4	3972	1	S75251
42	112	4.4	380	2	C83171
43	112	4.4	3198	2	A43426
44	111	4.4	1276	2	T18526
45	111	4.4	2508	2	S61441

ALIGNMENTS

RESULT	1
1000	1 1 00 8) large subunit (imported) - Sinorhizobium

alcohol dehydrogenase (acceptor) /
Sinorhizobium meliloti
 24-Aug-2001 #text_change 14-Sep-2001

C; Date: 24-Aug-2001 #sequence=

C; Accession: E95863

Author: Buhrmester, J.; Chain, P.; Vorholter, F.J.; et al.

Rifkin, Acad. Sci. U.S.A. 98, 3603-3607 (2001).
Proc. Natl. Acad. Sci. U.S.A. 98, 11481-11483 (2001).

A; Title: The complete works of William Shakespeare
A; Reference number: A95842; MUID: 21396508; PMID: 11101111

A; Status: preliminary
type: DNA

A: Molecular type: *Yersinia enterocolitica* O:3,4
A: Medication: 1-601 <KUR>
A: Cross-references: GB: A1591985; PIDN: CAC48573.1; PID: g15140045; GSPDB: G60040
A: Experimental source: strain 1021, megalaplasmid pSymB
A: Experimental source: strain 1021, megalaplasmid pSymB
A: Experimental source: S.R.: Phihler, A.: Abola, P.; Ampe, F.; Barloy-Hubb
A: Experimental source: S.R.: Phihler, N.A.; Fisher, R.

Rigalbert, F.; Finan, T.M.; Long, C.; Drenth, J.; Jones, P.W.; Jones, T.

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; L.; Hyman, A.M.; Science 293, 668-672, 2001
Weidenholter, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh

hebaullt, p.; vanuendert, ...
A;Title: The composite genome of the legume symbiont
A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation
C;Genetics:
A:Gene: Smb20173

A; Genome: plasmid
C; keywords: oxidoreductase

Query Match	16.9%	Score 427;	DB 2;	Length 601;
Local Similarity	24.9%	Pred. No. 1.2e-23;		
		202;	Indels 176;	Gaps 22

Best Local	72;	MISLADCEES
Matches 149; Conservative		
LANPAGEMISTYGNQENRHSPLQTITTEWGOLOLVW	71	

[illegible]

QY 72 A-----RGM-----PGKVFVPLIH-----
|| : ||| |

UUCUUUVVTTPRPNTVALDISKDGLWIKYEKPÖBENVIPVMCC 126

Db	67	TESTGYLRGHESSPLVAGDMLV	107
		-----IVANGVIVAGST-----	
R9	---	GVMYLAN---	186

```

      11 11 1
127 DTVNRGVAIADNKIFLHQADTVVALDAKTGKIWSYKNGDATKGETINAIWVAVKNDK
      -----PRAGEGDETW 145

```

QY 108 --COYSPPGCG--FVSGHDSATGEELMKNFLI
|| : : : ||| :
|| : : : ||| :
108 -NCTGCCCFEVRGHNTAYSMADGKVLMRGSMGPDDTLIDPEKTTHLGKRPVGKDSDGLTWTW 246

Db 18 VGISGGEVYNN
OY 146 GNDYEARMMTG---AWGQITYPVTLVHXSTAVGPASETORGIPIGGILXGINRFAVR 202

```

Db 247 EGD---QWKGIGTTCMGWSTYDEENLVYTGPNSTNPTQR--PGDNM-SMTTFARD 300
QY 203 PGTGEIWRHQTLPDMMDOECFFEMMTNVDVOPSTMEGLQINMNAATGRRYLVGV 262
Db 301 VDTGMAMKLYQMTPHDEMIDVGVNEMILLTEQDIDGK-----PRKLTJHF 344
QY 263 PCKTGMWQMDATGTEFL-----WARDTNYQWME-----SIDENGIVT 301
Db 345 D-RNGGCTYMDRTGELLVAEKYDPTVMATEVYMDKSKYGRPOVVAQYSTEQNG--- 400
QY 302 VNEDAILKELDVEYDCEPTFLGGRDWPSSALNPDGSGTYEPLNNVCYDMAVDOETSM 361
Db 401 --EDTWT-----GVCFAAGTQDQPPAAYSPKTELFYPTNHVCMDEFFRVSYTAG 452
QY 362 VYNTSNVTKLPKPKDM---IGRDAIDISTRTLMSEVERAANYSPVLTSGGVLENGCT 418
Db 453 PYVGATLSMPP-KDSHGGMGNFIANDNEKIKMSLEPFSVMSGALATAGDVFEYGLT 511
QY 419 DRYFRALSOETGETIMOTRLATVAASGATSYEDGMQYVATAG-----GVSYGSLNS 472
Db 512 EGYLKAADAATGKELYRFPKTPSGVIGNVMTTARESKQYVAVLSGVGGMAGIGLAQLTN 570

```

RESULT 2

```

S52317
quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
C:Species: Comamonas testosteroni
C:Date: 08-May-1995 #sequence, revision 21-Jul-1995 #text_change 02-Jun-2000
R:Accession: S62366; S62373; S65908; S52317
R:Stoovogel, J.; Krayveld, D.E.; van Stuks, C.A.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase
A:Reference number: S62366; MID:96184549
A:Accession: S62366
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STOI>
A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
A:Accession: S62373
A:Status: Preliminary
A:Molecule type: Protein
R:de Jong, G.A.H.; Geerlof, A.; Stoovogel, J.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purification
A:Reference number: S65908; MID:95324580
A:Molecule type: protein
A:Residues: 32-50; 'X', 52-54; 477-483; 'X', 485-490 <DED>
A:Experimental source: ATCC 15667
C:Genetics: qhdh
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; quinoprotein
F:1-31/Domain: signal sequence status predicted <SIG>
F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 status experimental <KMZ>

```

```

Query Match 16.7%; Score 422.5; DB 2; Length 708;
Best Local Similarity 23.5%; Pred. No. 3.3e-23;
Matches 139; Conservative 79; Mismatches 222; Indels 151; Gaps 18;

QY 3 PSLTMSAGALALAPAFQVTPVT-----DELLANPP-AGEWISTGONGEN 50
Db 12 PGRMWLLAACIG-SAAFAQTGSPAQAQAAAVORVGDGFIRANAFRPPDPTIGVDYAE 69
QY 51 YHSLPLOTTEENYGOLOLWARGMOPK-YOVTPLIHDGVVYLA----- 94
Db 70 TRYSLLDINANAKDGLAMSTNISTRTGVBATPVVDGIMYSASWSVVAIDTRTGN 129
QY 95 -----NIYANG----- 100

```

```

Db 130 RMTYDPOIDRSTGFKGCCVVRGVALNKGVYVAGMDGRLLALDAATGKEVWHONTPE 189
QY 101 -----VYASTCOYSPFGCCFVSGHDSATGELMRNFIT----- 135
Db 190 GOKGSLITTAGPAPVFKGKVIYIKRAEYGVG-YITVADMETGERKMFSPVGPSPK 248
QY 136 --RAGEGDEGTWQNDYEAARMGTG-----WQITVDDVTLVHYGSTAVGPASETORCTP 188
Db 249 EDESKMAARTW--DPSGKMEAGGCGTMDSMFDELMVTMVGTVGNGSPMSHKVRSK 306
QY 189 GGTLYGTNTRFARPDTEIYWRHQTLPDMMDOECFFEMMTNVDVOPSTMEGLQIN 248
Db 307 GGDNLVLAIVLADPDITKYMNTYQEPGDMMDYTSTOMILADIKI----- 353
QY 249 PNAATGERVRLTVPCCKTGMWQMDATGTEFLAARDTNTONNIESIDENG-IYTVNEDAI 307
Db 354 --AKRPRKVLHAP-RNGEFFVLDRTNCKFTSAKNFVPMVMSAGYDKHKKPIGI---AA 406
QY 308 LKELDVEYDCEPTFLGGRDWPSSALNPDGSGTYEPLNNVCYDMAVDO-EFT----- 358
Db 407 ARDSKRPQDAVPGRPYGANNHMSFNPOTGLVLYLPAQVAVNLMDKKMEFNQAGPKPO 466
QY 359 SMDYNTSNVTKL-PRCKMIGRIDAIDISTRTLMSEVERAANYSPVLTSGGVLENGCT 417
Db 467 SGTGMNTAKFNEAPKSKPFGRLAMPVQAKAAMSEVHSPWNGGTLTLAGNVVFOGT 526
QY 418 TDRFRALSOETGETIMOTRLATVAASGATSYEDGMQYVATAG-GGVSYSY 467
Db 527 ADRRLVAVYHAATGEEKLMEAPGTGVVAAPSTYVNDGRQYVAVVAGWGVYG 577

```

RESULT 3

```

A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence, revision 07-Apr-1994 #text_change 20-Jun-2000
R:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MID:94042848
A:Accession: A49340
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:9517067; PIDN:BA440252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

```

```

Query Match 16.6%; Score 421; DB 2; Length 742;
Best Local Similarity 24.7%; Pred. No. 4.6e-23;
Matches 150; Conservative 72; Mismatches 205; Indels 180; Gaps 20;

QY 11 AGALALLAPAFQVTPVYDELL--ANPPAGEWISTGONGENRHSPLQITTEENYGOLO 68
Db 23 AAALPYAAVPARADGGONTGEALITHADHPENLWSYGRYSQRYSPLOQIRSNVGDJK 82
QY 69 LVNARQMGQPKVQ-VIPLIHGCVYLT-----ANT----- 96
Db 83 LAMVYLLDTRNGQDEATPLVVDGIMVATTNMSKMEALDAATGKLMLQYDPKVGNIADKGC 142
QY 97 ----- 96
Db 143 CPTVNRGAGYNGKVFNGTDPGLVAAADATGKRWVNTIPADASLGKORSTYVDCAVR 202
QY 97 VANGYIVAGTCQYSPGCC--FVSGHDSATGELMRNFIRAGEED----- 142
Db 203 VAKGLVYLGN--GSEFGARGFVSADFAGTGLKWRFTYVNNKNEEDRAVADVNLMSKA 260
QY 143 -ETGNDYEAARMWT-----GAMGQITVDPYTNLVHYGSTAVGPASETORCTPGCTLY--- 193
Db 261 YKTMGP--KGAVNRQGGGGGTWDSLVYDPSDLTY---LAVG-----NGSPWNKRYKASE 309

```

[illegible]

```

RESULT      5
A:14270
Alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
A:Accession: S14270
R:Ramaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishida,
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membra
A:Reference number: S14270; MIMD:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:g216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane predicted <SIG>
F.1-35/Domain: signal sequence #status predicted <MAT>
F.36-738/Product: alcohol dehydrogenase 72K chain #status predicted

```

```

Query Match          16.1%:  Score 407.5:  DB 2:  length 738:
Best Local Similarity 23.7%:  Pred. No. 4,5e-22:
Matches 140:  Conservative 68:  Mismatches 219:  Indels 163:  Gaps 14:

10 SAGALALAAPAFQAQVTPYDE-----LLANPAGEMISGQNOENYRHSPLTOTTT 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
17 TACTICALISGYATMASADGGAGGAEALIHADHGNMNMGTGRYSQKSPILDQINR 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 ENNGOLQVWMAAGMQPKVOY-TPLIHGVML-----A 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
77 SNGNKLKMYLDDLTNRGEGTPLYDGMATTATNMSSMKAVDAATGKLLMSYDRVPG 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 NI----- 96
11 NIADKGCCDTVNGAIYNGKVFYGFEDGRLLALDARTGKLWMSYNTIIPPEALGKORSY 196
137 NIADKGCCDTVNGAIYNGKVFYGFEDGRLLALDARTGKLWMSYNTIIPPEALGKORSY 142
97 -----VANGVIAGSTCQISPPGC-FVSGHDSATGELMRYTFIPRAGEED----- 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 TYDGAFLRAAGRVILGN-GGSEFGARFYAFALAEKGKVDNRFYVPRKNEPDAASDS 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 -----ETWQNDYEAAMWTA-----KQITLYDPVTNLVLYGSTAVGASFOR 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
255 VLMNRAYQTFMS-----PTGATFROGGGGTVMDSIVYDPAVDLVLYGNGNSPWNRYKR

```

QY 186 GTPGGTLYGTNTREAVRPPDGEIWMRQTLPRDNDDECTEMAVTNDVQSTMEGLQ 245
 Db 308 SEGKGNLFLGSIYALKRGTETGTYWHEFTPEMDQWDFPSQINTLDP----- 357
 QY 246 SINPMAATGERRVLTVGPKCTGTMMQFDEATGEFFLMADINQOMIESID-ENGIVTVE 304
 Db 358 -----NGTRHVIYHARKNGFYIIDAKTGEFISGRKNVYVMAASGDPKTRPIYNP 410
 QY 305 DALIKELDYVDCPTFLGSRDMPAALNPDGSIYFPLNNVCYDMKAAVDOEFT-SMDVY 363
 Db 411 DALYTLTGKEMWGIPIGDLGHNPAAMAFSPRTGLVYIPAAQVPEFLTYNQVGSFPHDSW 470
 QY 364 NIS---NVTKLPPG-----KMGIRDAIDISTGRFLMSVERAANYSPLVSTGGV 412
 Db 471 NGLDMNKVGIPIPSPEAKQAFVVDLKGWIVAMPPOKQAMRVHDKGWMNGSILATGDL 530
 QY 413 LFNGGTDYFRALSOEFTGLMOTRLATVASGQAISEYEDQMVAIAGG 462
 Db 531 LEQGLANEFHAYDATNGSDFHFADSGIIMPVYTYLANGKQYVAEYV 580
 RESULT 6
 J00706
 N:Altemate names: methanol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
 C:Species: Methylobacterium extorquens
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
 R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
 A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 moxF and moxJ genes
 A:Reference number: J00706; MUID:90337342
 A:Accession: J00706
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-626 <AND>
 A:Cross-references: GB:M31108; NID:9150017; PIDN:AAA25380.1; PID:9150018
 R:Nunn, D.N.; Day, D.; Anthony, C.
 Biochem. J. 260, 857-862, 1989
 A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
 A:Reference number: S04644; MUID:89350892
 A:Accession: S07908
 A:Molecule type: Protein
 A:Residues: 28-50, 'XX', 53 <MUN>
 A:Note: the source is designated as Methylobacterium extorquens AM1
 C:Comment: This enzyme oxidizes methanol to formaldehyde.
 C:Genetics:
 A:Gene: moxF
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match 15.7% Score 396.5; DB 2; Length 626;
 Best Local Similarity 23.9%; Pred. No. 2,3e-21;
 Matches 141; Conservative 84; Mismatches 205; Indels 159; Gaps 22;
 QY 10 SAGALALLA-AP-AFAQVTPVDELANPPAGE-WISYGOENRHRSHLQITTEVNGQ 66
 Db 7 SVSALAMIALAPALSSGAYANDKVELSKSDNWMVPRKNTDSNNESDLQIKNGVKKQ 66
 QY 67 LQLVW--ARGQPGKVVQVPLIHGQVYL----- 93
 Db 67 LRPATYFSTGLNGH-EGAPLVVDGKMYIHTSPNNKTRALGLDPPGTLWODKPKQNPAA 125
 QY 94 -----ANIVA-----NGYIVAGSTC 108
 Db 126 RAVACCDLVNRGLAYWPGDGKTPALILKTQLDGNVAALNAETGETWYKVENSDIKVSTL 185
 QY 109 QYSPF-----GCVSCHDSATGEELMRYFI-----P 135
 Db 186 TIAPYVKKQVVIIGSSGAEIYVRYTAYDVKTGEQVRAIATGPKDLLASDENIKNP 245

QY 136 RAGEE--GDETWGNDYARMRTGA---WQIYDPVTLVHGSTANGPASETORSTPGC 190
 Db 246 HYGGKGLCTGTWSD---AMKIGGGTNNMGWYAYDGGTLLIFGTGNPAPWNETMR--PGD 300
 QY 191 TLVGTNTRFAVRPPDGEIWMRQTLPRDNDDECTEMAVTNDVQSTMEGLQ 245
 Db 301 NKW-TMTIFGRDADTEAKFGYQKTPHDEM-----YAGVNVMLSEQD----- 344
 QY 251 AATGERRVLTVGPKCTGTMMQFDEATGEFFLMAR---DYNQOMIESIDENGIVTVE 304
 Db 345 -KDKARKLLTPHNDNGVLYTLDRDGLVSAKLLDT--VVFVSVLTKTGQPRDPEY 401
 QY 308 LKELD-VEYDVCPTFLGSRDMPAALNPDGSIYFPLNNVCYDMKAAVDOEFT-SMDVY 363
 Db 402 GTRMDHLAKDLCFSAWYHNOGHDSDYDPKRELFFMGINHCMDPEPMLPYRAGOFFVGA 461
 QY 367 NVTKLPPGK-----DWGRIDAIDISTGRFLMSVERAANYSPLVSTGGVLFNGGTR 420
 Db 462 TLNMPGPKGDRQYBGLQIKAYNAITGDKYKEMERFAVWGTMATAGDLVFTGLDG 521
 QY 421 YFRALSOEFTGLMOTRLATVASGQAISEYEDQMVAIAGG 462
 Db 522 YLKRSDTDGLLWKFIPSGAIGYPMYTHKQYVAT-----YGVG 565
 RESULT 7
 S68591
 N:Altemate names: methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylobacterium methylobacter (str
 C:Species: Methylobacterium methylobacter
 A:Title: Methylobacterium methylobacter
 C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
 R:Kia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.
 J. Mol. Biol. 259, 480-501, 1996
 A:Title: Determination of the gene sequence and the three-dimensional structure at 2.
 A:Reference number: S68591; MUID:96256524
 A:Accession: S68591
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-573 <XIA>
 A:Cross-references: EMBL:U41040; NID:91127819; PIDN:AAA83765.1; PID:91127820
 A:Note: The authors did not translate the codons for residues 1-2
 C:Keywords: oxidoreductase

Query Match 14.2% Score 359.5; DB 2; Length 573;
 Best Local Similarity 23.0%; Pred. No. 1,1e-16;
 Matches 127; Conservative 75; Mismatches 187; Indels 153; Gaps 20;
 QY 38 AGEWISTGQNDENRHRSHLQITTEVNGQLVW--ARGM----- 75
 Db 13 AGAMPATGGYGGQSHNSPLAQINKSVKAVKAAVSFGVNLNGHGAPLYIDMMYVNSA 72
 QY 76 -----OPGK-----VQVPLIHGQVNLAVIY--ANGVIVA- 104
 Db 73 FPNNTYALNLPBKIKWQHCKPRKODASTKAVMCCDYVDRSLAGAQYKQAQNGHLAL 132
 QY 105 -----GSTCYSPF-----GCVSCHDSATGEELM 129
 Db 133 DAKTGIMNEVEVCDPKVGSSTLQAPFVAKDYLMGSCAELGVGAIVNAPDLETKGLK 192
 QY 130 R-----NYFIPRAGE--EGDETWGNDYARMRTGA---WQIYDPVTL 168
 Db 193 RAEATGSDSVRLAKDPSANPHYGFGLGTITWGD--AMKIGGGTNNMGWYAYDPPKLN 249
 QY 169 LVHGSTAVGPASTGCTGTPG-----TLVGTNTRFAVRPPDGEIWMRQTLPRDNDDE 223
 Db 250 LEFYGSQNPAPWNETMR--PGDNKWTMTIMGRD-----DTGAAMKGYQKTPHDEM 301
 QY 224 CTFEMVNTNDVQSTMEGLQSIQINPNAATGERRVLTVGPKCTGTMMQFDEATGEFFLMAR 283
 Db 302 GYVQNVYLTJ-----QPVN-----GKMPPLSHIDRNGILYTLNRENGNLIVAE 344

F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93,95/Binding site: ubiquitinone (Arg, Asp) #status predicted
 F:466/Active site: Asp #status predicted

Query Match 11.8%; Score 299.5; DB 1; Length 796;
 Best Local Similarity 21.3%; Pred. No. 4,7e-14;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

```

QY 6 LMAS-----AGALALAAAPAOVTVTDELLANPAGEMISYGOENRHSPLTQ 58
DB 135 LTMAGFNDPQRLNGTISADAPPAEA-ISPVDQ-----DMPAIGRNDGQRFSPKQ 185
QY 59 ITTENVGOLQVLA-----RGMOPGV--OVPLIHDSVMTLNI-----96
DB 186 IMADNVHNLKAWVFRGDKVQPNPDGEITNEVTFIKVGDLYLCTAHQRLFALDAASGK 245
QY 97 -----VANGVYA-----104
DB 246 EKMHYDPELKTNSFOHYTCRGVSYHEAKAETASPEVMADCPRIILPVNDGRLAIINAE 305
QY 105 -GSTCO-----YSP-----FCGFV 117
DB 306 NGKLCETFPANKVNLQSNMPTKPGLYEPTSPIITDKTIYVAGSVTDFSTRETSGLV 365
QY 118 SGHDSATGEELMNRNYFIPIRA-----GEGDETWGNDYEARMWTGANGOLITDPTVNLV 416
DB 366 RGFVNTGELLMA--FDGAKDPNAIPSDHEFTFENS-----PNSWAPAAIADALDLY 416
QY 171 HGSSTAIVPASETRGRTGCGTLTGNTREAVRPDGTGEIYVHROTLPBNDMDCETFEMNV 230
DB 417 YLPMGVTTPLDINGKNETPEQERYASSI-LALNATTKLMSYOTVHHDLMDM-----468
QY 231 TNDVOPSTEMEGLOSINPNAATGERVLTGVPCKTGMQFDETEGL-----280
DB 469 --LPAOPT-----LADITVN--GQKVPYIAPAKTGNTFVLDNRNCELVELVAPAEKVPVQ 518
QY 281 WARDTNYQNMIESIDENGI-----VTNEDAILKEI--DVEYD-----316
DB 519 GAAKGDVYTPPOFSELSFRPTKDLGADMMGATMFDQLVCRVMPHQMRTEGTFPPSEQ 578
QY 317 ---VCPFTLGGDWPSSALNP-----D 335
DB 579 GTLVFPGNLGFMFEGISVDPNREVAIANPMLPVSCLKLPKPGCNPEQPKDAKGTGE 638
QY 336 SGI-----YFIPLNNVCYDMAVDQETSMQVNTSNTVKLPKPKDMIGRIDALDIST 388
DB 639 SGIOPOYGVPRGYTLN-----PFLS-----PFLGPCQKOPAMGYISALDLKT 679
QY 389 GRTLSVERAANYS-----PV-----LSTGGVLFNGCT--DRYFALSOE 428
DB 680 NEVVMKKRIGTPDQSMPEPMPVPVFPNMGMPMLGGPISTAGNVLFIAATADNYLRAVYMS 739
QY 429 TGETLMQTRLATVASGQA--ISYEVDGMQYVAI-AGGCVSYGS 468
DB 740 NGEKLMQGRLP--AGQATPMTEYVNGKQYVVISAGHGSGFGT 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0505)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA833551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECS0128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.8%; Score 299.5; DB 2; Length 796;
 Best Local Similarity 21.3%; Pred. No. 4,7e-14;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

```

QY 6 LMAS-----AGALALAAAPAOVTVTDELLANPAGEMISYGOENRHSPLTQ 58
DB 135 LTMAGFNDPQRLNGTISADAPPAEA-ISPVDQ-----DMPAIGRNDGQRFSPKQ 185
QY 59 ITTENVGOLQVLA-----RGMOPGV--OVPLIHDSVMTLNI-----96
DB 186 IMADNVHNLKAWVFRGDKVQPNPDGEITNEVTFIKVGDLYLCTAHQRLFALDAASGK 245
QY 97 -----VANGVYA-----104
DB 246 EKMHYDPELKTNSFOHYTCRGVSYHEAKAETASPEVMADCPRIILPVNDGRLAIINAE 305
QY 105 -GSTCO-----YSP-----FCGFV 117
DB 306 NGKLCETFPANKVNLQSNMPTKPGLYEPTSPIITDKTIYVAGSVTDFSTRETSGLV 365
QY 118 SGHDSATGEELMNRNYFIPIRA-----GEGDETWGNDYEARMWTGANGOLITDPTVNLV 416
DB 366 RGFVNTGELLMA--FDGAKDPNAIPSDHEFTFENS-----PNSWAPAAIADALDLY 416
QY 171 HGSSTAIVPASETRGRTGCGTLTGNTREAVRPDGTGEIYVHROTLPBNDMDCETFEMNV 230
DB 417 YLPMGVTTPLDINGKNETPEQERYASSI-LALNATTKLMSYOTVHHDLMDM-----468
QY 231 TNDVOPSTEMEGLOSINPNAATGERVLTGVPCKTGMQFDETEGL-----280
DB 469 --LPAOPT-----LADITVN--GQKVPYIAPAKTGNTFVLDNRNCELVELVAPAEKVPVQ 518
QY 281 WARDTNYQNMIESIDENGI-----VTNEDAILKEI--DVEYD-----316
DB 519 GAAKGDVYTPPOFSELSFRPTKDLGADMMGATMFDQLVCRVMPHQMRTEGTFPPSEQ 578
QY 317 ---VCPFTLGGDWPSSALNP-----D 335
DB 579 GTLVFPGNLGFMFEGISVDPNREVAIANPMLPVSCLKLPKPGCNPEQPKDAKGTGE 638
QY 336 SGI-----YFIPLNNVCYDMAVDQETSMQVNTSNTVKLPKPKDMIGRIDALDIST 388
DB 639 SGIOPOYGVPRGYTLN-----PFLS-----PFLGPCQKOPAMGYISALDLKT 679
QY 389 GRTLSVERAANYS-----PV-----LSTGGVLFNGCT--DRYFALSOE 428
DB 680 NEVVMKKRIGTPDQSMPEPMPVPVFPNMGMPMLGGPISTAGNVLFIAATADNYLRAVYMS 739
QY 429 TGETLMQTRLATVASGQA--ISYEVDGMQYVAI-AGGCVSYGS 468
DB 740 NGEKLMQGRLP--AGQATPMTEYVNGKQYVVISAGHGSGFGT 780

```

RESULT 11
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potlowski, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85495
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <STO>
A:Cross-references: GB:AE005174; NID:912512839; PIDN:AG54428.1; GSPDB:GN00145; OMGP:201
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gsd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.8%; Score 299.5; DB 2; Length 796;
Best Local Similarity 21.3%; Pred. No. 4.7e-14;
Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;
DB 6 LMAS-----AGALALLAFAAQTPTVDELLANPAGWISYQONQENYHSPLTQ 58
135 LTMAGFNDPOEINGTLRADATPA-A-ISPVAQD-----DMPAYGRNDEGGRSPLKQ 185
QY 59 ITTENVGOLQVMA-----RGMOPGV--QVTPLEHGVYLANI----- 96
DB 186 ITADNVHOLKEAWVFRGTGDKQPNDEITNEVTPLEKVDLYLCTAHQRLPALDAASGK 245
QY 97 -----VANQVIVA----- 104
DB 246 EKMHPDLKTNESFOHVTGCVSYHEAKAETASPEVMADEPRRIILPVNDGRLFAVNAE 305
QY 105 -GSTCQ-----YSP-----FGCFV 117
DB 306 NGKLCETFAKNGVNLQSNMPTKRGLEYEPTSPITITKTIYVAGSYTDFSTRETSYI 365
QY 118 SGHDSATGEELMRNYTPRA-----GEEDETWGDYEAARMKMGQITVDPTNLY 170
DB 366 RGDVNTVGLLMA--FDPGAKDPNAPDEDETFEENS-----PNSWAPADAIDAKIDL 416
QY 171 HGSTAVGASSETQKGTGCTGNTFAVAPDTEIWRHQLPRDNDDECTEEMV 230
DB 417 YLPMGVTPDIMGNTPEQERYASSI-LALNATGKLAWSYQVTHDMDMD----- 468
QY 231 TNDVOPSTEMEGLOSINPNAATGERRVLTGVPCKTGTMGDFETGEF----- 280
DB 469 --LPAQPT-----LADITVA--GOKPVITAPARTGNIFFVLDNRNGELVPAPEKVPQ 518
QY 281 WARDTNTQNMIESIDENGI-----VTNEDALIKEL--DVEYD----- 316
DB 519 GAAGADYVTPQPSSEISFPTKDLGADMGATMFDOLVCVMEHQRYEGIFTPPESEQ 578
QY 317 ---VCPFLGGRDWPSSAALNP-----D 335
DB 579 GTLVFPGMLGMEFGISYDPNREVAIANPMALPFVSKLIPRGPNPMEQPKDAKGTGE 638
QY 336 SGI-----YFTPLNNVCYDMAVADQEFMSDVYNTSNVTKLPPGKDMIGRIDAIDIST 388
DB 639 SGIOPOYGVPPGYTLN-----PFLS-----PFLGPKQAPMAGYISALDKT 679
QY 389 GTLMSVERAANYS-----PV-----LSTGGVLYFNCGT--DREFRALISOE 428
DB 680 NEVYAKKRIKTRPOSMPEPMPVPPFNNGMPLGSPITAGNVLFIAATADNLTIRAYNMS 739
QY 429 TGETLMQRLATVAGQA--ISYEVDGMQYVAI--AGGVSYS 468
DB 740 NGEKIMQGRLP--AGGQATPMTEYNGKQYVVISAGHSGSFGT 780

RESULT 12

AG0523
glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhimurium
A:Note: this species has also been called Salmonella typhimurium
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG0523
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connelton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.: Moutle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Agonaedis
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0523
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:916501455; GSPDB:GN00176
C:Genetics:
A:Gene: STY0191
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.7%; Score 295.5; DB 2; Length 796;
Best Local Similarity 21.5%; Pred. No. 9.4e-14;
Matches 148; Conservative 66; Mismatches 207; Indels 267; Gaps 29;
DB 6 LMAS-----AGALALLAFAAQTPTVDELLANPAGWISYQONQENYHSPLTQ 58
135 LTMAGFNDPOEINGTLRADATPA-ATSSIADE-----DMPAYGRNDEGGRSPLKQ 185
QY 59 ITTENVGOLQVMA-----RGMOPGV--QVTPLEHGVYLANI----- 85
DB 186 ITADNVHOLKEAWVFRGTGDKQPNDEITNEVTPLEKVDLYLCTAHQRLPALDAASGK 245
QY 86 -----IDGVYLANI----- 95
DB 246 EKMHPDLKTNESFOHVTGCVSYHEAKAETASPEVMADEPRRIILPVNDGRLFAVNAE 305
QY 96 -----VANQVIVA-----FGCFV 117
DB 306 NGKLCETFAKNGVNLQSNMPTKRGLEYEPTSPITITKTIYVAGSYTDFSTRETSYI 365
QY 118 SGHDSATGEELMRNYTPRA-----GEEDETWGDYEAARMKMGQITVDPTNLY 170
DB 366 RGDVNTVGLLMA--FDPGAKDPNAPDEDETFEENS-----PNSWAPADAIDAKIDL 416
QY 171 HGSTAVGASSETQKGTGCTGNTFAVAPDTEIWRHQLPRDNDDECTEEMV 230
DB 417 YLPMGVTPDIMGNTPEQERYASSI-LALNATGKLAWSYQVTHDMDMD----- 468
QY 231 TNDVOPSTEMEGLOSINPNAATGERRVLTGVPCKTGTMGDFETGEF----- 280
DB 469 --LPAQPT-----LADITVA--GOKPVITAPARTGNIFFVLDNRNGELVPAPEKVPQ 518
QY 281 WARDTNTQNMIESIDENGI-----VTNEDALIKEL--DVEYD----- 316
DB 519 GAAGADYVTPQPSSEISFPTKDLGADMGATMFDOLVCVMEHQRYEGIFTPPESEQ 578
QY 317 ---VCPFLGGRDWPSSAALNP-----D 335
DB 579 GTLVFPGMLGMEFGISYDPNREVAIANPMALPFVSKLIPRGPNPMEQPKDAKGTGE 638
QY 336 SGI-----YFTPLNNVCYDMAVADQEFMSDVYNTSNVTKLPPGKDMIGRIDAIDIST 388
DB 639 SGIOPOYGVPPGYTLN-----PFLS-----PFLGPKQAPMAGYISALDKT 679
QY 389 GTLMSVERAANYS-----PV-----LSTGGVLYFNCGT--DREFRALISOE 428
DB 680 NEVYAKKRIKTRPOSMPEPMPVPPFNNGMPLGSPITAGNVLFIAATADNLTIRAYNMS 739
QY 429 TGETLMQRLATVAGQA--ISYEVDGMQYVAI--AGGVSYS 468
DB 740 NGEKIMQGRLP--AGGQATPMTEYNGKQYVVISAGHSGSFGT 780

RESULT 13

S00943
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoaceticus
C:Species: Acinetobacter calcoaceticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S00943
 R:Cleaton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A:Title: Nucleotide sequence of the gene coding for quonoprotein glucose dehydrogenase
 A:Reference number: S00943; MID:88289368
 A:Accession: S00943
 A:Molecule type: DNA
 A:Residues: 1-801 <CLE>
 A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712
 A:Experimental source: strain LMD 79.41
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:33-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 11.48; Score 289; DB 1; Length 801;
 Best Local Similarity 20.28; Pred. No. 2,9e-13;
 Matches 140; Conservative 76; Mismatches 174; Indels 304; Gaps 30;

20 PAFQVTVTDELNPAGEWISYQNGENRHSPLTQTTEENVGOLQVARGMOPGK 79
 152 PETAQAVGVAE-----SMPAYGRTQAGVRSPLKQINDQNVKDLKAVMT--LRIGD 202
 80 V-----QVTPPL----- 85
 203 LKTNDGSETNQVTPPIKIGNNFICTAHQOLAIIDPATGKEKWFDPKLTGKSFQHLT 262
 86 -----HDGVNYLAN----- 95
 263 CRGVMYDANNTTFFATSIQSSKSSSTQCPKRVFVAVNADTGKACTDFGNG 322
 96 -----IVANGYVAGS--TCQYS-----PGCVSHDSATGEEL 128
 323 QVNLOEFMPYAPGYNPTSPGIYVIGSVTNDTSNKEPSG-VIRQYDVNTGKL 381
 129 WRNFIPTAGE-----EGDETGWNDYERAWMTGAMGOITDPVNLVHGSGTAVGASE 182
 382 W--VFTDGAADPNAMPGEETFEVHNS-----PNAMAPLADAKLDIV-IYPTGV----- 427
 183 TQGTGEGTLYG-----TNTFAVRPDGTGEIYVRRHOTLPRWDQECTFEEMVYN 232
 428 ---GTF--DIMGDRTELKERANSMLAINASTGKLWNFQTHHDLMDVDSQPSLAD 482
 233 VDVQPTSTEGSLQSTNP--NATGERRYLTGP-----CKTG--TMMQDAET 276
 483 INKAGQIVPAIYLTGKNAFVLDLR--NGQPIVPTKEPVQYAKRGQTKGEFYSKY 540
 277 GEF-----LMA-----RDTNYQMTESIDENGIVTVNEDAIL 308
 541 QFESDLNAPDKLTDKMDMGATMLDQLMCRVSEKRLNMDGIVTTPSENGTL----- 592
 309 KELDEIVDYCTFLGGRDMPAALNPS-----GIYFI----- 341
 593 -----VFPGNLGVFEWGMGSVNPDRQAVANNPVIGLFPVSLLPADENRAQTAKGAGT 644
 342 -----PLNNVCYDMAVDQFTSMDDYNTNSVTKLPFGKMDIGRIDAIDISTGRTLM--- 393
 645 EQGVQPMYGVY-----GVEISAFSLPLGLPCCKPANGVYAGVDLKTHEVWAKR 694
 394 -----SEVERAANYSPVLSTGGGVLFNGST--DRIYRALSOETGETLMQTR 437
 695 IGTIRDSLPHLFDLPVAKTKIGVPGIGSGISTAGNVMEVATQDNLYLRANVTNGKLMWEAR 754
 438 LAYVASQA--ISYEDVGMQYVAI--AGGGSVSYGS 468

Db 755 LP--AGGQATPMTYELNKGQYVIMAGHSGSFGT 786

RESULT 14

OPKEX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda
 C:Species: Gluconobacter oxydans
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Cleaton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 MOL. Gen. Genet. 229, 206-212, 1991
 A:Title: A single amino acid substitution changes the substrate specificity of quhop
 A:Reference number: S17716; MID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CLE>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, 'A', 214-808 <GO>
 A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417
 C:Genetics:
 A:Gene: gdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:9-28/Domain: transmembrane #status predicted <TM1>
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:122-138/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:470/Active site: Asp #status predicted

Query Match 9.9%; Score 251.5; DB 1; Length 808;
 Best Local Similarity 18.8%; Pred. No. 1.7e-10;
 Matches 132; Conservative 77; Mismatches 212; Indels 281; Gaps 24;

13 ALALLAPAPQVTPVYDEL-----NPPAGEWISYQNGENRHSPLTQIT 61
 128 AVLAFASLTDPDHDHISGELPTQIANASPADPDNPVASEMHAIGRTQAGDWSPLQINA 187
 62 ENVGQLOLV-----ARGMOPGKV--QVTPLIHDGVNYLAN----- 96
 188 TVNSNLKVAHHTKDMANSNDPEQOTNEATPIEFNNLTLYMCSLHOKLFAVDGATGNVK 247
 97 -----VANGYIV----- 103
 248 VYDPKLIQINPFOHLTCRGVSFHETPANAMDSQNPATPCAKSILIPVNDGRLVEVDAD 307
 104 AGSTC-----OYSPFGCV----- 117
 308 TGTGCGFNNNEIDLRPVNPQYTTGGYEPTSPVITDKLIIANSATIDNGSVKQASGA 367
 118 -SGHDSATGEELW---RNVEIPRAGEGDETWGNDYERAWMTGAMGOITDPVNLVHY 172
 368 TQAFPVYTGKRWYVDASNPDPNQLPDESHPVHPNSPMSIVS-----SYDANLILYI 422
 173 GSTAVGPASEFQGRFP---GSTLYGNTFR---AVRPDGEIYVRRHOTLPRDNDQEC 224
 423 PMGV-----GTPQMGQDRTKQSERAPGIVALLNDGTGLAMFYGYTHHDLMDL 473
 225 TFEWVNTVNDQPTSTEGSLQSTNPNAATGERRYLTGVCCKTKGTMQDAETG--EFLMAR 283
 474 PSEPDLVDVYQKQSTLVPAIYA-----PTKTGDIFFVLDRTGKEIYVAP 517

```

Dh      443  PDMWGNNRPGAEKFSAGLYALDNTGKLRMYQYTHHDLMDM-----VGSQPTL 493
Qy      240  EMBGLQSIINPAATGERRVILTVGCKTGMQFDETG-----                277
Dh      494  -----IDLKADQVKKALI-APTKQGSILYLDLRDDGPIVPIRENPAPQAGVEGDHRA 545
Qy      278  -----EFILWA-----RDNYQNMIESIDENGITYTNEDAI 307
Dh      546  PTOGARDLNLRLPLTERDMWGSSPFIOMCLIQFSLAYEGQVPPSEQSL----- 598
Qy      308  LKELDEVYVCPTEFLGGDMWPSAALNPDSCGTYIPLINANCYDMAVDOEPT-----SMQVYN 364
Dh      599  -----IYGNVGYNWGWGSVDYDROIETTSNRYAFVSQWVPDKYPPSSCKREGE 649
Qy      365  TSNVTK-----LPGKGMIGRIDAIDISTGRTLSVERAA-NISPV 405
Dh      650  TSGVQNPITGAPYAVIMHPMSPITGLPQAGASWGVAGVIDLTLLAKVVOHKNGKTSMDNPT 709
Qy      406  -----LSTGGVLENGGT-DRFRALSGEGETLMQRLATVASGQA--IS 448

```

```
OY      449 YE-VDGMQIYAI-AGGSVSYS 468  
        |  |  |  |  |  |  |  |  
Db      768 YTGKDGRQYLIIVAGGHSFGT 789
```

Search completed: May 24, 2002, 10:21:03
Job time: 341 sec

Fri May 24 11:27:37 2002

wallick-934-95.pep.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:22 ; Search time 28.06 Seconds
(without alignments)
651.305 Million cell updates/sec

Title: WALICK-934-95.PEP
Perfect score: 2530
Sequence: 1 MKPTSLWASAGALALAAP.....GMOYVALAGGVSYSGSLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	16.3	739	1 DHET_ACEU	Q44002 acetabacter
2	409	16.2	742	1 DHET_ACEU	P18278 acetabacter
3	407.5	16.1	728	1 DHET_ACEU	P28036 acetabacter
4	401.5	15.9	757	1 DHET_GLUOX	O05542 gluconobact
5	400	15.8	631	1 DHET_PARDE	P12293 paracoccus
6	396.5	15.7	626	1 DHET_METEX	P16027 methylobact
7	396.5	15.6	600	1 DHET_PARDE	P29968 paracoccus
8	395.5	15.6	626	1 DHET_METEX	P16027 methylobact
9	357	14.1	571	1 DHET_METEX	P38339 methylobact
10	356.5	14.1	571	1 DHET_METEX	P38339 methylobact
11	299.5	11.8	796	1 DHG_ACEU	P05465 acetabacter
12	289	11.4	801	1 DHG_ACEU	P27175 gluconobact
13	251.5	9.9	808	1 DHG_GLUOX	O59086 acetabacter
14	209.5	8.3	790	1 DHG_ACEU	O9478 acetabacter
15	189.5	7.5	827	1 AFSK_STRCO	O60282 methanococc
16	119	4.7	799	1 AFSK_STRCO	O60282 methanococc
17	118.5	4.7	799	1 AFSK_STRCO	O60282 methanococc
18	115.5	4.6	353	1 YXAL_BACSU	P42111 bacillus su
19	115.5	4.6	353	1 YXAL_BACSU	P42111 bacillus su
20	115	4.5	443	1 PORO_PSEAE	O99614 mus musculu
21	114	4.5	796	1 COPP_SCHPO	P32722 pseudomonas
22	111	4.4	1276	1 SCAP_CRIGR	O42937 schizosacch
23	110.5	4.4	3317	1 CADN_RAT	P58365 ratius norv
24	108.5	4.3	3354	1 CADN_RAT	O9h251 homo sapien
25	108	4.3	466	1 MM08_RAT	O88766 ratius norv
26	105.5	4.2	593	1 SCAP_STRSP	P19909 streptococc
27	104.5	4.1	1279	1 SCAP_HUMAN	O12770 homo sapien
28	102	4.0	1645	1 PLMN_PIG	P06867 sus scrofa
29	102	4.0	1645	1 PLMN_PIG	P06867 sus scrofa
30	101	4.0	1012	1 YFGL_FCOLI	P77774 escherichia
31	101	4.0	1012	1 YFGL_FCOLI	P77774 escherichia
32	100	4.0	465	1 MM08_MOUSE	O70138 mus musculu
33	100	4.0	500	1 VLI_HPV6A	P03100 human papil

34	99.5	3.9	807	1 AFSK_STRGR	P54742 streptomyc
35	99	3.9	752	1 GCP2_RAT	P70627 r glutamate
36	99	3.9	1365	1 GTPS_STRPO	P29336 streptococ
37	98.5	3.9	1016	1 PMPH_CHLIR	O84880 chlamydia t
38	98.5	3.9	1256	1 FINC_CHICK	P11722 gallus gall
39	98.5	3.9	566	1 AMY_STRGR	P30270 streptomyc
40	97.5	3.9	501	1 VLI_HPV11	P04012 human papil
41	97.5	3.9	746	1 FEPA_FCOLI	P05825 escherichia
42	97	3.8	649	1 ACES_DROME	P07140 drosophila
43	97	3.8	1355	1 RP0D_ANASP	P22705 anabaena sp
44	97	3.8	2265	1 FINC_BOVIN	P07589 bos taurus
45	96	3.8	566	1 AMY_STRLM	P09794 streptomyc

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	739 AA.
DHET_ACEU				
ID	DHET_ACEU	Q44002: 007952:		
AC	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).			
GN	ADH.			
OS	Acetobacter europaeus.			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;			
OC	Glucanobacter.			
OX	NCBI_TaxID=33995;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	THURNEI C.A.K. / DSM 6160;			
RA	Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.			
RL	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.			
CC	-1- COPACITOR: POO AND HEME (BY SIMILARITY). CYTOCHROME, AND TWO			
CC	-1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, DEHYDROGENASE			
CC	-1- SMALLER UNKNOWN SUBUNIT) THAT FORMS THE ALCOHOL DEHYDROGENASE			
CC	COMPLEX (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC			
CC	SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch)			
CC	-----			
CC	EMBL: X82894; CAA58066.1; -			
CC	EMBL: X09480; CAA70688.1; -			
DR	HSP: O92477; LFIQ.			
DR	InterPro: IPR001479; Bac_POO.			
DR	InterPro: IPR002372; Bac_POO_repeat.			
DR	InterPro: IPR000345; CytC_heme_bind.			
DR	Pfam: PF01011; Bacterial_POO_6; 6.			
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C; 1.			
KW	Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.			
FT	SIGNAL	35		
FT	CHAIN	36	739	
FT	BINDING	651	651	HEME (COVALENT) (BY SIMILARITY).
FT	BINDING	654	654	HEME (COVALENT) (BY SIMILARITY).
FT	METAL	655	655	IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ	SEQUENCE	739 AA;	80944 MW;	E681BB237ACB91F4 CRC64;

Query Match 16.3%; Score 412; DB 1; Length 739;

Best Local Similarity 24.0%; Pred. No. 7.4e-24;
Matches 140; Conservative 68; Mismatches 226; Indels 150; Gaps 14;

```

OY 10 SAGALLALAPAFQVTPYVDE-----LANPAGEMISYGOENYRHSPLTOTTT 61
DB 17 TAGTICAMISGATATMSADGGGATGEMAIHADHPGMMMYGRTYSQRRSPLDQINR 76
OY 62 ENVOLOLWARKGQPKQYV-TPLIHGVMYL-----A 94
DB 77 SNVGNLKLAWYLDLDFNRGQEGTPLYIDGMATINMSKAKVDAATGKLLMSYDPVPV 136
OY 95 NI-----
DB 137 NIADKGGCDTVNRGAAYWNGKYVFGTGDRLTALDAKTGLWMSYNTIPPEALGKORSY 196
OY 97 -VANGYVAGSTQYSPFC--FVSGHDSATGEELMRNYFLPRAGEEDFTWGN 147
DB 197 TVDAPRIKAGVITIGN--GSEFGARGVAFALFCKVDKMRFFTPAKRNPBDHTASD 254
OY 148 D-----YEARMTGA-----NGOITYDPTNLVHYGSTAVGASSETGRTPGGT 191
DB 255 SYLMNKAYQWSPTCAMTRGCGGCTWDSIVYDPAVLVYGNGSPWNYKRSCKD 314
OY 192 LGTITRPAVRPDGEIYVHRQTLPRDNWQECTFEKAVTVNDVQSTMEGLQSTNPA 251
DB 315 NLFDSIYALRPETGEYVWFQETPMQMDPTSYQOIMTLDLPINET----- 362
OY 252 ATGERVLTGVPCKTGTWMOEDAGTEFLMAROTNQNMTESID-ENGITVNEALIKE 310
DB 363 -RHVIVAP-KNGFFIITDAKTGEISGKNYVYVNASGLDPTGPRIPNPALVTL 417
OY 311 LDVEVDVCPITLGRDMPALNPDGSIYFIPLANNYDMAVDEFT-SMNYVTS--- 366
DB 418 TGKMTYIGIPDGLGHNFAAFSPKGTGLVYIPAOQVPELYTNOVGFTPHDPSMWLGDM 477
OY 367 NVTKLPPG-----KDMGRIDALDISTGRTLSVEAAANISPVLTSGGVLENGGT 418
DB 478 NKVGIPIPSPAKQAFVKDLKGVIVAMPDQKQAEAMRVYDHKGPMNGILATGDLFGOLA 537
OY 419 DRYERALSQETGETLMOTRLATVASQAISYEVDGMQYVAIAG 462
DB 538 NGEFHATDNTGSDLFHFAADSGIAPVYTLANGKQYVAVEVG 581

```

RESULT 2
DHET_ACEAC STANDARD: PRT: 742 AA.
AC P18278:
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_Taxid=435;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 36-50.
RX MEDLINE=9525070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RA Yano K.;
RT Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti.;
RL J. Bacteriol. 171:3115-3122(1989).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=9528964; PubMed=7772016;
RA Cozier G.E., Giles T.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens".

```

RL Biochem. J. 308:375-379(1995).
CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -I- COFACTOR: POO AND HEME.
CC -I- SUBUNIT: TETRAMER OF NON IDENTICAL CHAINS (DEHYDROGENASE,
CC CYTOCHROME AND TWO SMALLER UNKNOWN SUBUNITS).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: D90004; BAA14058.1; -.
DR PIR: J50326; J50326.
DR HSSP: Q924J7; 1FLG.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_6.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR OXidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 742
FT ACT_SITE 343 343
FT BINDING 649 649
FT BINDING 652 652
FT METAL 653 653
SQ SEQUENCE 742 AA; 81521 MW; 9C6C9268DAB825A CRC64;

```

Query Match 16.2%; Score 409; DB 1; Length 742;
Best Local Similarity 24.2%; Pred. No. 1.3e-23;
Matches 149; Conservative 78; Mismatches 192; Indels 196; Gaps 22;

```

OY 11 AGALALLAPAFQVTPYVDELT--ANPAGEMISYGOENYRHSPLTOTTTENQQLQ 68
DB 23 AALLPYAAVAPARADGGGNGEATIHADHPENWLTSGRTYSQRRSPLDQINRSNVGDLK 82
OY 69 LVNARGMQPKQYV-TPLIHGVMYL-----ANI----- 96
DB 83 LLGYTLIDTRNGGEATPLVVDGIMATNTNMSKEMALDAATGKLLMOYDPKVPGLADKGC 142
OY 97 -----
DB 143 CDTVNRAGYWGNGKVFWGTGFDGRVLAADAKTGKRVMAVNTIPADASLGKORSTYDGAVR 202
OY 97 VANGYVAGSTQYSPFC--FVSGHDSATGEELMRNYFLPRAGEED----- 142
DB 203 VAKGLVITIGN--GSAEFGARGVSAFAETGKLMKRFYTPVNNKKNPNHNASDNLIMNKA 260
OY 143 -ETWGNDEYKAMNT-----GAMGOITYDPTNLVHYGSTAVGASSETGRTGCTGLY--- 193
DB 261 YKTWGP--KGAMVYRGGGGVTWDSLVYDPSDLT---LAVG-----NGSPWNYKRSY 309
OY 194 -GTN-----TRFAVRPDTGELVYHRQTLPRDNWQECTFEKAVTVNDVQSTMEGLQSI 247
DB 310 GIGSLNLFGLSIVALKPETGEYVWFQETPMQMDQDYSYQOIMTLDMPK----- 358
OY 248 NPNATGE-RRVLTVGPCKTGTWMOEDAGTEFLMAROTNQNMTESID-----EN 297
DB 359 -----SEMRHVIYVAP-KNGFFIYVDAKTGEISGKNYVYVNASGLDPTGPRIPNP 411
OY 298 GIYTVNEDALIKELDEYVDCPTFLGGRDMPALNPDGSIYFIPLANNVC----- 348
DB 412 GLVTLNG-----KFWYGI-PGPIGAHNFAAFSPKGTGLVYIPAOQVPELYTNOVG 463

```


QY 349 -----DMMAYDOFTSMYNTSNVTKLPKPKDMIGRIDADISNGRTLMSEVERAANYSP 404
 DB 464 KPHADSMNNGVLDWTKNGLPDPE-ARTAYIKDLHGMFLAMPVMEYWKIDHKGPNNNG 522
 QY 405 VLSGGGVLENGSTDTYFRLALSOETGELTMOGRLATVASGOAISYEVDQMYYA----- 458
 DB 523 ILANGGLLEFQGLANGEEFAYDATNGSDLKRDAGSGLIAPPMYTSVNGKQYVAVEVGW 582
 QY 459 -----IAGGVSYSGS 469
 DB 583 GIYPSMGVGRSG 597

RESULT 3

ID DHET ACEPO STANDARD; PRT; 738 AA.
 AC P28036;

DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.

OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 NC Acetobacter.
 NCBI_TaxID=439;

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=NB11028;

RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;

RT "Cloning and sequencing of the gene cluster encoding two subunits of
 membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL);
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.

CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE

CC COMPLEX
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@sib-sid.ch](mailto:license@sib-sib.ch)).
 CC -----

CC EMBL: D00635; BAA00528.1;
 CC PIR: S14270; S14270.

DR HSP: Q9Z4J7; JPLG.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR OXIDOREDUCTASE: POQ; Heme; Periplasmic; Membrane; signal.
 KW OXIDOREDUCTASE.
 FT CHAIN 1 35 POTENTIAL.
 FT BINDING 36 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 738 AA; 80840 MW; 1E2B6D7BCD92AF6 CRC64;

Query Match 16.1%; Score 407.5; DB 1; Length 738;
 Best Local Similarity 23.7%; Pred. No. 1.6e-23;
 Matches 140; Conservative 68; Mismatches 219; Indels 163; Gaps 14;

QY 10 SAGALLAIAFAFAOYTPVDE-----LAPNPAEMISYGNQENYRHSPLQITTT 61
 DB 17 TAGTICALALISGYATMASADDGCGATGEALITHADHPGMMMYGRYSQORRSPIDQINR 76
 QY 62 ENVGQLOLVARAGQPOCKVQV--EPLIHGVYVL-----A 94
 DB 77 SNVGNLKLMTLIDLTNRGQCGPLVDVYVYATVMSMKAVDAATGKLMSYDPRVPG 136
 QY 95 NI----- 96
 DB 137 NIADGCCDVTNRGAAYNGKRYEGTEPDGRLIALDAKTGLVSWVTIIPPEALGKQRSY 196
 QY 97 -----VANGYIVAGSTCOYSPFGC--FVSGHDSATGEELMNNYFIPRAGEGD----- 142
 DB 197 TVDGAIRLAKGVITIGN--GGSEFGARGFVSAFDAETGKVDMEFFVPPNKNPDASDS 254
 QY 143 -----ETWGNDEARWMTGA-----WQIYDPVTLVHGSTAVGPASETOR 185
 DB 255 VLNKRAYQWMS-----PTGAMTRQGGGGTVMDSIVYDPAADVLYGVNGSPMNYKRA 307
 QY 186 GTPGGTLYGTNRFAVRPDTGEIVRHOITLPRDMWDQCTPEMNTVNDVOPSTEMGLQ 245
 DB 308 SEGKGNLFLGSLVALKPEETGEVYHFPQETPDMDQDPFSDQIIMTLDLPI----- 357
 QY 246 SIPPNATGERRYLVGPKGTGMQFPAETGEFLMADTYVONMIESID-ENGIVTVNE 304
 DB 358 -----NGETRHVYHARKNGFFYIIDAKTGEISGNKYVYVWASGLDKRTPIYNP 410
 QY 305 DAIKELDVEYDQPTFLGGRDPSAALNPDGIFYFIPLNNVCYDMAVDOEFT-SMDYV 363
 DB 411 DALYTLGKEMVYIGPDLGSHNFAAASFPGKGLVYPAQVYFLTYNQVGGFTPHPSW 470
 QY 364 NTS---NYTKLPD-----KDMIGRIDADISGRTLMSEVERAANYSPVLTGGGV 412
 DB 471 NLGIDMKNVGIIPDSPEAKQAFVVDLKGIVAMPQKQAEAMRWDRHKGPNNGIATGDDL 530
 QY 413 LFNSTDRFRLALSOETGELTMOGRLATVASGOAISYEVDQMYYAIAAG 462
 DB 531 LFOGLANGEFHAYDATNGSDLFHPSADSGIAPPVYILANGKQYVAVEVG 580

RESULT 4

ID DHET GLUOX STANDARD; PRT; 757 AA.

AC 005542;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH
 DE subunit I).

GN ADHA.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 NC Gluconobacter.

NCBI_TaxID=442;
 RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
 RC STRAIN=IFO 12528;
 RX MEDLINE=97208225; PubMed=9055427;

RA Kondo K., Horinouchi S.;
 RT "Characterization of the genes encoding the three-component membrane-
 RT bound alcohol dehydrogenase from Gluconobacter suboxydans and their
 RT expression in Acetobacter pasteurianus.";
 RL Appl. Environ. Microbiol. 63:1131-1138(1997).

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL)
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL)
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86375; BAA19753.1; -
DR HSSP: Q924J7; IFLG.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam: PFO1011; Bacterial_POQ.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 6.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
KW SIGNAL 1 34
FT CHAIN 1 34
FT MOD_RES 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT ACT_SITE 342 342 PYRROLIDONE CARBOXYLIC ACID.
FT BINDING 653 653 BASE (POTENTIAL).
FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;

Query Match 15.9%; Score 401.5; DB 1; Length 757;
Best Local Similarity 23.0%; Pred. No. 4.9e-23;
Matches 141; Conservative 79; Mismatches 217; Indels 175; Gaps 18;

QY 7 LMSAGALALIAA--PAFAQV--TPVDELNPPAGEISGQNOENYRSPLOTOT 61
DB 16 LLSCAALALFAAFAVFAQEDTGFATISDSNGHP-GMLISGHSYSORSPLODQNT 74
QY 62 ENVGQQLVVARGMQPKYOV--TPLIHDGVML--
DB 75 ENVGKLLAHVYLDLTNRGQETPLVNGVMATNWSKMLDAATGKLLMSYDPKPG 134
QY 95 NIYANG--
DB 135 NIADRGCCDTVSRAAYWNGKVFEGTFDGRLLALDAKTGLVWSVYTIPEAOLGHORSY 194
QY 101 ----VIVAGSTCOYSPFGCEVSGHDSATGELMRNFFPRAGECD----- 142
DB 195 TVDGARIRAKGVLLGNGAEGFARG-FVSAFDAETSKLDRFPTVENPEKPDGAASD 253
QY 143 -----ETWGNDYEARWMTG--AMQIYDPTNLVHGSTAVGASSETOGTGGTL 192
DB 254 ILMKRAYPTWKGKAMKQGGGCTWDSLVDPYDVLVGLVNGSGSPWNNKFFSECKGN 313
QY 193 YGNTREAVRPTGEIVRHOITLPRDNMDQCTEFEMAVTVNDVOPSTEMEGLSINPMA 252
DB 314 LFLGSIVAINPDTGKYVHFQETPMDEDDYTSVQOITLMPV----- 356
QY 253 TGE-RRVLGVPCKTGTWQDAETGEFLMAROTNQNMTESIDE-NGIVYNEAALIKE 310
DB 357 NGEHNVIVAP-KNGFFYIIDAKTGFTGCKPYTEENMANGIDPVYGRPNVPPALMTL 415
QY 311 LDVEYDVCPTFLGGRDPSAALNDGIVPIPLNNV-----CYDMAVDDDEF 357
DB 416 TGRPLUGLPGELGHNFAAAYSPKTYIVYIPAOQIPLLYDQGGKFAKHYHAWMLGLDM 475
QY 358 TSMQVNTSVYTKLPCKDMIGRIDALIDISTGRL-WSTERAANYS-----PVL 407
DB 476 NKIGLFDNDPBEHAAKKD-----LKVYKGTVAWDEKMAPAFTINHKGPMNGGLA 529

QY 408 TGGVLENGTDRYPRALSOETGETLMOTRLATVASGASISYEVDMQYAT----- 459
DB 530 TAGNVTFQGLANGFEHAYDATNGNDLVSFPAQSAIINPPVYITANGKQYAAVEGNGGTY 589
QY 460 --AGGVSYGSG 469
DB 590 PLYGSGVARTSG 601

RESULT 5
DHML_PARDE
ID DHML_PARDE STANDARD; PRT; 631 AA.
AC P12293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus
OX NCBI_Taxid:266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RX MEDLINE-87307969; PubMed-3114231;
RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stouthamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
RT structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC ON METHANOL: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17339; AAA88366.1; -
DR HSSP: P38539; AAAH.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam: PFO1011; Bacterial_POQ_7.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR Oxidoreductase; POQ; Signal; Methanol utilization; periplasmic.
KW SIGNAL 1 32
FT CHAIN 1 32
FT DISULFID 135 136 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT_SITE 335 335 BASE (POTENTIAL).
SQ SEQUENCE 631 AA; 69799 MW; 0934DC93FEC5730B CRC64;

Query Match 15.8%; Score 400; DB 1; Length 631;
Best Local Similarity 24.3%; Pred. No. 5e-23;
Matches 146; Conservative 76; Mismatches 200; Indels 176; Gaps 23;

QY 4 TSLMASGALALAL-APAFQVTPVDEL--LANPAGEISGQNOENYRSPLOTOT 60
DB 12 SSLAAVAMGLAVLTAPATA-----NDQVLELAKDPA-NWVVTGSDYNAQYSEKTDIN 65
QY 61 TENVGQQLVVA----- 72

QY 308 LKELD-VEYDVPFELGSDWPSAALNPDSCITFIPILNNVCYDMAVDEFTSMQVNTS 366
 DB 402 GTRMDHLAKDICPSAMGTHNQHDSDIPKRELFEMGINCHMDWEPFMLPYRAGGFFPGA 461
 QY 367 NTKLIPCK-----DMIGRIDAIDISTGRTLSVERAANSPVLSTGGVLFNGGTRD 420
 DB 462 TLNMPGPGRDRONYEGISQIKAYNAITGDYKKEKERRAVWGTMATAGDLVEFGTLDG 521
 QY 421 YFRLASOETGELTMOPTLATVASGOAISYVDGMQVYALVAGGVSGS 469
 DB 522 YLKARDSDTGDLMKRKIRISGALIGPMTYHKGTOYVAI-----YGVG 565

RESULT 7

XOXF_PARDE STANDARD; PRT; 600 AA.
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RA SEQUENCE FROM N.A.
 RL Harms N.;
 RP Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 RC STRAIN=PD 1235;
 RX MEDLINE=92041583; PubMed=1657873;
 RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
 RT Stoutamer A.H.;
 RT "Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c551 of Paracoccus denitrificans and characterization of
 RT the mutant strain."
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1- COFACTOR: POQ (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.sib.ch).
 CC CC
 DR EMBL: U34346; AAC44555.1;
 DR EMBL: M75583; AAA25574.1;
 DR PIR: A41378; A41378.
 DR HSP: P38539; 4AAH.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR Oxidoreductase; POQ; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 600
 FT ACT SITE 318 318
 FT ACCT SITE 600 AA; 65159 MM; DCA996F1BCC5A3CE CRC64.
 SQ SEQUENCE

Query Match 15.6%; Score 395.5; DB 1; Length 600;
 Best Local Similarity 24.7%; Pred. No. 1e-22;
 Matches 148; Conservative 75; Mismatches 204; Indels 173; Gaps 24;

QY 10 SAGLALALAPAPAYQVPTDELANP-----PAGEWISYGVQENYRSPILTQTTE 62
 DB 6 NGACLALMSGTAA-----LANEQRAGRDRQAPQWALQMGDYANTRYSLTDQINKD 56

QY 63 NGQGLQLVN--ARGNQPKGVQVPTLIHDGVYKLANIVANGV----- 101
 DB 57 NKKDLRAVMTFTSTGLVRH--EGSPFLVIGDWYVHTPPNNVFALDINDKILMRYEPOQ 115
 QY 102 -----IYAGSTC-----QYSPGCFVSGH-----DSATGELMRNF--IPRAGE--- 139
 DB 116 DENVIAVMSCTPYRGLSYAD--GMILIGQADITVVALDATSQGVKSTIGPGIGELT 174
 QY 140 -----EGDEW-----GNDYEA----- 151
 DB 175 ATVPVAKDKVLVGISGEYGVGRKMTALNLTDSSEAKMASTGPDELLVDPEITHLGK 234
 QY 152 -----RMWTGA-----WGQITDPTNINLVHGSTAVGASETGTGTGTLGK 194
 DB 235 PIGADSLNSWEGDQWQIGGTTWGFESYDPLNLVYCTGNPSTWNSQR--FGDNKW- 291
 QY 195 TNRFAVRPPTGELIWRHQTLPDNDQCTFEMWVTN--VDVQSTMEGLQSIINPAA 252
 DB 292 SMTIMARDADTGMAKWFYQWTPHDEMDYGVNEMILITQVDDGQ----- 335
 QY 253 TGERRVLTGVPCKGTGMQFPAETGELFWARDNYQNMIE-----SIDENGIYTVNEDAI 307
 DB 336 --ERKLTHFD--RNGLAYTMDRETGELLVAE--KYDPVNMNTGVDMDPS--ETGRPAV 389
 QY 308 LKELDVEYD-----VCPFLGSDWPSAALNPDSCITFIPILNNVCYDMAVDEFTS 359
 DB 390 VAEYSTAONGEDENNTGVCYPALGTRKQDQAPAFSPKRNILYVPTNHCMDYEPFRVAYTA 449
 QY 360 MDVY--NTSNVTKLPKPGDMIGRIDAIDISTGRTLSVERAANSPVLSTGGVLFNGG 417
 DB 450 GQPYVATISMTAPADNSHGGNGFTAMHNTGTIKMSVPQFSWGSALATAGDVYFT 509
 QY 418 TDRYRALSOETGELTMOPTLATVASGOAISYVDGMQVYALVAGGVSGS 472
 DB 510 LRGYLKPVDAQTGELYKFTKPSGIQNWYTERHGKQYVIGILSGVAGIGLAAGLTN 569

RESULT 8

DHML_METOR STANDARD; PRT; 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit I precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXE.
 OS Methylobacterium organophilum XX.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 29-43.
 RC STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
 RX MEDLINE=89008094; PubMed=2459109;
 RA Machlin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 RT methylobacterium organophilum XX methanol dehydrogenase structural
 RT gene."
 RL J. Bacteriol. 170:4739-4747(1988).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Query Match 15.6%; Score 394.5; DR 1; Length 626;
 Best Local Similarity 23.9%; Pred. No. 1.3e-22; Indels 159; Gaps 22;
 Matches 141; Conservative 85; Mismatches 204;

10 SAGALALLA-AP-APAOYPTVDELLANPAGE-WISYONQENYRHSPLTQITTEVGO 66
 7 SYVALMLALAPALSAVAYANDKVELSKDDMMVMPGKNYSDNNYSELKOVNKNVQ 66
 67 LQLVW--ARGMOPGVQVYPLIHGVNVL----- 93
 67 LRAMVFSGLNGLH-EGAPLVYDGRMYVHSFPNNTPALDLDPRHLMQDKPKNPAA 125
 94 -----ANIVA-----NGVVASTC 108
 126 RAVACDVLNRGLAYWPGDGKTPALILKQIDRHYVALNAETGETVWKVENSIDIKVSTL 185
 109 QVSPF-----GC-----FVSGHDSATGEIMRNFYI-----P 135
 186 TIAYVYVCKVYIIGSSGAEVLGVRGLTAVDVTGGVRAVATGDKDLLADDENVKNA 245
 136 RAGE--GDETWGNDYFARMYGA--WQIYDPTVNLVHSTAVGPASTQRTGPTGG 190
 246 HYGQGLGATATWED--AMKIGGTNWGVAYDPSTNLIFGTGNPAPWNETMR--PGD 300
 191 TLYGTNFFAVRPDTGETVWRHQTLPDNDNDOCTEMEVNTVNDVQSTEMEGLOSINPN 250
 301 NKW-TMTIFGRDADTGEAKFQYKTPHDEWDYAGVNVAM-----PSQKX----- 344
 251 AATGERVLTGVPCKTGMFOFDETEGFLMAR---DTYONKIESIDENGIYVNEDEAI 307
 345 -KDGKTRKLLHPDRNGIYVTLDRDGLALVANSKLDPT--VNVFKTVDLKTGPVADPEY 401
 308 LKELD-VEYDVCPTFLGRDMPSAALNDGSIYFIPANNVCYMMVANDQETSMDEVYNTS 366
 402 GPFMDLADVCPASAMGHNOGHSYDPRKLEFPMGINHLCOMDEPMLPYRAGQGFVGA 461
 367 NVTKLPRCK-----DMIGRIDAIDISTGRILMSVERAANYSEVLSGGGVTFNGCTR 420
 462 TLNMYGPRGDRONTGEGLOIKAYVALTGSYKMEKMERFAVWGGLATADLVFYESTLDG 521
 421 YFRALSOETGETIMOTRLATVATVSAQAISEYVDGMQVYIAAGSVYSVG 466
 522 YLAKRSDTGDLLMKFKIPSGAIGYPMYTHKGTQYVAI-----YGVG 565

RESULT 9
 EXAA_PSEAE STANDARD; PRT: 623 AA.
 AC 094J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (QEDH).

GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PAOI;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Smith K.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
 RA opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schodet M., Goerisch H.;
 RA "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RA oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RA the gene encoding cytochrome c550 and an adjacent acetaldehyde
 RA dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-202023376; PubMed-10736230;
 RA Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RA "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RA Pseudomonas aeruginosa: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 RT -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC -1- ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC "THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announce/>
 CC OR SEND AN EMAIL TO license@isb-sib.ch).

FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;

Query Match 14.18; Score 357; DB 1; Length 623;
 Best Local Similarity 21.7%; Pred. No. 9.5e-20;
 Matches 130; Conservative 87; Mismatches 231; Indels 152; Gaps 19;

QY 3 PTSLMASAGALAL--LAAPFAQVTPYDELAL--PPAGEWISYQONENRHSPLT 57
 DB 9 PAGLLRSLHCLAFVALGSAALAKADYTMEDIANDKTGDVLYQMGTHAQRWSPK 68
 QY 58 QITTEENGOQLYVWANGM---QPGKVYVPLIHGVAVL----- 93
 DB 69 QVANDNVEFLTPAMSYSPFDEKORQO-ESQALVSDGVITYVASYRFLADAKTKRLMT 127
 QY 94 -----ANIYA-----NGVIV-----AGST 107
 DB 128 YNRLPDDIRPCDDVYNNRGAALYCKVFFGLDASVVALNKNCKVYKKRKFADHGAGYT 187
 QY 108 CQVSP-----FCGF--VSGHDSATGEEELMRNYFIR----- 136
 DB 188 MGAFLPYKDGKTKVLLIHSSGDEFGVGRFLFARDPDTGEELIWRPFEGHGRSLNGK 247
 QY 137 ---AGEEGDETWGNDYEA-----RMTGA---WGOITDPVYVNLVHGSTAVGPASET 183
 DB 248 DSTVTGDKVAKSPDDKRSPTGKVESHSKGAPWOSASAPDABTNIIYAGNPGPWNW 307
 QY 184 ORGTGPG---TLGTNTRFAVRPDTGEIVRHOQLPDMDOCTEFEMMTVAVDQV 237
 DB 308 APTAKGNPHDYDLY-TSGGVGVDPSSGEVKWFYQHTNDAMDSGNNELVLFYKAKD 366
 QY 238 STEMEGLQSIINP-----AATGERVILTVPECKTGTW--QFDAETGEFLMARDNTQNM 290
 DB 367 GKIVKATAHADRNGFEVYVDRNSKGLQNAFPEVDNITWASHIDLKTG-----RPERBGQ 421
 QY 291 IESIDENGITVVEDALIKELDVEYDCTFLGGRWPSAALNPDGTYFPIPLNNVCYDM 350
 DB 422 RPLPEPG-----OKHGRAVEVSPPLGKNWPMAYSDQITGLFYPAHMKEDY 471
 QY 351 MAVDOFTSMDEVYNTSNVTKLPCKMIGRAIDISIGRLMSVERAANYSPVLSTGG 410
 DB 472 WTEEVSYTKSAVLGKGFRIKMYDDHVSGLRAMPVSGKVMWKEKHLPLMAGVLTAG 531
 QY 411 GVLFGNGTDRYRALSOETGETLMQTRLATVYASGSAISYEVDGQYVAILAGGVSGSL 470
 DB 532 NLVFTGTGDTGKAPDAKSGKELMKFQTSGISVSPITWEDDQGYLGV---VGTGAV 588

RESULT 10

ID DHM1_METME STANDARD; PRT: 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha subunit) (MDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; Beta subdivision; Methylophilus group;
 OX NCBI_TaxID-17;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA MEDLINE-9405969; PubMed-8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RT "The active site structure of the calcium-containing quinoprotein
 methanol dehydrogenase.";
 RL Biochemistry 32:12955-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).

RX MEDLINE-93054513; PubMed-1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution.";
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1 CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.
 CC -1 COFACTOR: TWO MOLECULES OF POQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1 SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1 SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1 SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 DR PDB: 4AAH, 08-DEC-96.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1, FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2, 1.
 KW Calciolum.
 KW Oxidoreductase; POQ; Methanol utilization; Periplasmic; 3D-structure;
 FT DISULFID 103 104
 FT DISULFID 379 408
 FT ACT_SITE 297 297
 FT HELIX 2 9
 FT STRAND 11 12
 FT STRAND 14 14
 FT STRAND 17 18
 FT STRAND 21 22
 FT STRAND 26 27
 FT STRAND 34 36
 FT STRAND 37 39
 FT STRAND 41 47
 FT STRAND 59 61
 FT STRAND 62 63
 FT STRAND 64 68
 FT STRAND 71 73
 FT STRAND 75 79
 FT STRAND 80 81
 FT STRAND 83 84
 FT STRAND 86 90
 FT STRAND 96 101
 FT STRAND 103 104
 FT STRAND 112 114
 FT STRAND 115 116
 FT STRAND 121 121
 FT STRAND 123 124
 FT STRAND 126 131
 FT STRAND 132 134
 FT STRAND 137 142
 FT STRAND 146 148
 FT STRAND 149 149
 FT STRAND 151 151
 FT STRAND 157 159
 FT STRAND 160 161
 FT STRAND 162 166
 FT STRAND 168 168
 FT STRAND 170 172
 FT STRAND 173 173
 FT STRAND 177 182
 FT STRAND 183 185
 FT STRAND 188 193
 FT STRAND 198 202
 FT STRAND 203 203
 FT STRAND 205 210
 FT STRAND 212 213
 FT STRAND 222 222
 FT STRAND 223 223
 FT STRAND 226 231
 FT STRAND 241 243
 FT STRAND 244 247

BASE (POTENTIAL).

FT STRAND 248 252
 FT HELIX 261 263
 FT TURN 269 270
 FT STRAND 273 278
 FT TURN 279 281
 FT STRAND 284 289
 FT STRAND 305 311
 FT TURN 312 313
 FT STRAND 314 322
 FT TURN 324 325
 FT STRAND 327 332
 FT TURN 333 335
 FT STRAND 338 343
 FT TURN 346 347
 FT STRAND 351 355
 FT TURN 356 359
 FT STRAND 360 363
 FT HELIX 365 367
 FT STRAND 370 372
 FT TURN 371 372
 FT STRAND 373 373
 FT TURN 375 378
 FT STRAND 382 383
 FT TURN 390 390
 FT STRAND 392 394
 FT TURN 395 398
 FT STRAND 399 405
 FT TURN 407 413
 FT STRAND 420 421
 FT TURN 427 433
 FT STRAND 435 436
 FT TURN 438 439
 FT STRAND 444 450
 FT TURN 451 454
 FT STRAND 455 462
 FT STRAND 469 469
 FT TURN 471 473
 FT STRAND 474 476
 FT TURN 477 481
 FT STRAND 483 484
 FT TURN 486 491
 FT STRAND 492 495
 FT TURN 496 502
 FT STRAND 512 516
 FT TURN 517 518
 FT STRAND 519 526
 FT TURN 530 533
 FT HELIX 534 538
 FT TURN 539 539
 FT TURN 543 545
 FT HELIX 546 552
 FT TURN 553 555
 FT HELIX 556 558
 FT TURN 559 559
 FT STRAND 565 570
 SQ SEQUENCE 571 AA: 62449 MW: 0BE94EA5A2DAB1E1 CRC64:

Query Match 14.1%; Score 356.5; DB 1; Length 571;
 Best Local Similarity 23.5%; Pred. No. 9.2e-20;
 Matches 130; Conservative 75; Mismatches 192; Indels 157; Gaps 22;

QY 38 AGEWISYGOQENYRHSPLTQTTENVGQLOLV- - - - -
 DB 11 AGAMPATGGYSSQHSPLAOLNKSNNVKNKAMSFSGVNLNGHGAAPVIGDMYVHSA 70
 QY 76 - - - - -
 DB 71 FENNTYALNLDPEKTIWQHKPKQDASTKAVKCCDVYDVGGLAYGAGQIVKQKQSHLLAL 130
 QY 105 - - - - -
 DB 131 DAKTKINMEVEVCDPKVGSITLTOAPFVAKDYVLKMGCSAGELGVGAVNAPDLTKTGELKW 190

QY 130 R-----NFIPIRAGE--EGDETCNDYEARWMTGA---MGQITYDPVTN 168
 DB 191 RAFATGSDSVRLAKDNPANPHYGQGLCTKTWED---AMKIGGTNMGWYAYDRLN 247
 QY 169 LVHGSTAVGPASEFORGTGPG-----TLVGTNRFAPRDPDTGELVWHPHLLPRDNDOE 223
 DB 248 LFYYSGNPAPNETMR--PGDNKWTMTINGRDL-----DTGMKAKWIOKTPHDEMFOFA 299
 QY 224 CTFEMWTVNDVOP--STEMEGLQS--INPNATGERRVLTGVPCKTGTMMQDAETGEFLW 281
 DB 300 GVNQAVLTLD--QPVNAKMTPLLSHIDRN-----GILYTLNRENGNLIIV 340
 QY 282 ARDTN--YOMMIESIDENGIVTVNEDAILKELDE--YVCTPFLGGRDMSALNPDGSIY 339
 DB 341 AEKVPAVNEFKVDLKTGTVPKDEFEATRDHKTINCPSAMGFHNGVDSYDPSRFL 400
 QY 340 FIPLNVCYDMAVDOEFTSMDEVNTSNVTKLP-----PGKMDICRIDAIDISTGRILNSV 395
 DB 401 YAGLNHICMDMEPFMLPYRAGQFVGATLAMPQPNQPTKEMGOIRAFDLITGKAKWK 460
 QY 396 ERANANSPLYLSTGGVLFNGGTDRTYFRALSOETGETLWOTRLATVASGOAISYEVDMQ 455
 DB 461 WEKFRAMGTLTYKGLVWYATLDGYLKALDNKDKELMFRMPSGGIGSPMTYSFKGQ 520
 QY 456 YVAIAGGVSTYSG 469
 DB 521 YI-----GSMYGVG 529

RESULT 11
 ID DHG_ECOLI STANDARD; PRT; 796 AA.
 AC P15877;
 DT 01-APR-1990 (rel. 14, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).
 GN GCD OR B0124.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12; PubMed=2228962;
 RA Cleton-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
 RT "Cloning, mapping, and sequencing of the gene encoding Escherichia
 coli quinoprotein glucose dehydrogenase.";
 RL J. Bacteriol. 172:6308-6315(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RA Yamada M., Asakura S., Saiter M.H., Jr., Yamada Y.;
 RT "Characterization of the gcd gene from Escherichia coli K-12 W3110
 and regulation of its expression.";
 RL J. Bacteriol. 175:568-571(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RA MEDLINE=94261430; PubMed=8202364;
 RT MEDLINE=94261430; PubMed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA MEDLINE=97426617; PubMed=9278503;
 RT MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12." ;
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=93286127; Pubmed=8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.,
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site." ;
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96128046; Pubmed=8554505;
 RA Cozier G.E., Anthony C.,
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *coli* modelled on that of methanol dehydrogenase from *Methylobacterium*
 RT *extorquens*." ;
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X51323; CA35706.1;
 DR EMBL: D12651; BA02174.1;
 DR EMBL: D26562; CAB20298.1;
 DR EMBL: AE000122; AAC73235.1;
 DR PIR: JY0107; JY0107.
 DR HSP: P38339; 4AAM.
 DR EcoGene: EG10369; gcd.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR ProSITE: PS00363; BACTERIAL_POO_1; 1.
 DR ProSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 37 40 PROBABLE.
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PERIPLASMIC (PROBABLE).
 FT DOMAIN 142 796 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT ACT_SITE 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TLSADAMP -> HLKRRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;

Query Match 11.88; Score 299.5; DB 1; Length 796;
 Best Local Similarity 21.38; Pred. No. 3.3e-15;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

OY 6 LMAS-----AGATALLAARFAQVTPYVDELLANPAGEWISYGOONENRHSPLQ 58
 DB 135 LTMAGFNDPEINGTSLADATPAEA-ISPVADO-----DMPAYRNOEGGRFSPLO 185
 OY 59 ITTENVOQLDLVA-----RGMPGRV--QVPLIHDMVMTLAN----- 96
 DB 186 IADNVHNLKEAMVFRFTGDKQPNPDGEITNEVPPIKVGDTLYLCTAHOFLPALDAAGK 245
 OY 97 -----YANGVYA----- 104
 DB 246 EKHYDEBELKTNESFOHTKRGVSTHEAKETAPEVMADCPRIILPVNDRLAINAE 305
 OY 105 -GSTCQ-----YSP-----FGCEV 117
 DB 306 NGKLETPANKGLVNLDSNMDPTKPGLYEPTSPIITDTIYAGSVYDNFSTRETSQVI 365
 OY 118 SGHDSATGEELMRVNFIPRA-----GEGDFTWGNDEYARMWTGANGQITDPVTLV 170
 DB 366 RGFVNTGELLMA--FDGAKDPNALPDSDEHTFENS-----PSWAPAYADAKLDLV 416
 OY 171 HYGSAVAPASRTORGTPGGTLYGTNTFRVAPRPDGEIYVRRHOTLPKRNMDQECTFEMAV 230
 DB 417 YLPMGVTPPIDIMGRTPEOERYASSI-LALNATTKLANSYOTVHHDLMDMD----- 468
 OY 231 TNVDVOSTEMEGLOSINPAATGERRVLTPCCTGTMOFDAETGEFL----- 280
 DB 469 -LPAQPT-----LADITVA--GOKVPVYAPAKTGNIFVLDNRNGLVVPAPKPVQ 518
 OY 281 WARTQNVQNNIESIDNGI-----VYVNEPALIKEL--DYVED----- 316
 DB 519 GAAGKGVLPPTQFSSLSFRPTKDLGADMMGAMFMPOLCRVWFHOMRGEITPPPSQ 578
 OY 317 ---VCFPLGGRDMPSAALNP-----D 335
 DB 579 GTLVFPGNIGMEWKGISVDNREYALANPALPEVSKLIPRGPNMPQPKDANGTGE 638
 OY 336 SGT-----YIPLANNVCYDMAVDOFTSNDVYNTSKLPPEKMDIGTIDIDIST 388
 DB 639 SGIQPOYGVVGYTLN-----PFLS-----PFLGPKOPAMGYISLIDLKT 679
 OY 389 GRTLMSVERAANYS-----PY-----LSTGGCVLENGGT-DRYFRALSOE 428
 DB 680 NEVYMKRRTIGTPDSMPFMPVPPVPMKMGPMIGPISAGVLEFIATADNYLRAVIMS 739
 OY 429 TGETIMORLATVASGA--ISTYVDKQYAI-AGGVSYS 468
 DB 740 NGKIMOGRLP--AGGQATPMYEVNGKQYVVISAGHSGSGT 780
 RESULT 12
 DHGA_ACICA STANDARD; PRT; 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH-A.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_Taxid=471;
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMD 79.41;
 RX MEDLINE=88289368; Pubmed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 RT dehydrogenase from *Acinetobacter calcoaceticus*."
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.

CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINOTOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 CC POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07235; CAA30222.1; -
 CC PIR: S00943; S00943.
 CC InterPro: IPR001479; Bac_POO_repeat.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
 CC Oxioreductase; POO; Transmembrane; Periplasmic; Signal.
 CC SIGNAL 1 33 POTENTIAL.
 CC CHAIN 34 801 GLUCOSE DEHYDROGENASE-A
 CC [PYRROLOQUINOLINE-QUINONE].
 CC FT TRANSMEM 39 55 POTENTIAL.
 CC FT TRANSMEM 59 79 POTENTIAL.
 CC FT TRANSMEM 94 108 POTENTIAL.
 CC FT TRANSMEM 119 138 POTENTIAL.
 CC FT ACT_SITE 471 471 BASE (POTENTIAL).
 CC FT SQUENCE 801 AA; 86956 MW; 2FAEA222B5450D8 CRC64;
 SO

Query Match 11.4%; Score 289; DB 1; Length 801;
 Best Local Similarity 20.2%; Pred. No. 2,1e-14;
 Matches 140; Conservative 76; Mismatches 174; Indels 304; Gaps 30;

OY 20 PAFQAVPVYDELLANPAGEWISYQONENRHSPLQITTEYNGQLYVWARKMPQK 79
 DB 152 PETQAVVGVAVF-----SDMPAYGRTQAGVRSPLKQINDQVNDKLVAVT--LRTGD 202
 OY 80 V-----QVPL-----
 DB 203 LKTNDGSETTNOVTPKIGNNMFCITAHQOLIAIDPAIGKEKMRDPKLTGKTSFOHLL 262
 OY 86 -----IHDSVMTLAN----- 95
 DB 263 CRGVMYDANNTTFAFATLSQSKSSSTQCPKRVFPVNDRLVAVANADTGKACTDFGONG 322
 OY 96 -----IVANGVYAGS--TCQYS---PFCGFVSGHDSATGEEL 128
 DB 323 QVNLQFMPYAPPGYNTSPGIYGVSTVYAGSVTQNSKKESS--VINGYDVNTGKLL 381
 OY 129 WRYVTFPRAGE-----EGDETGWGNDYEARMWTGAGQOITVPTNLVHGSTANGPASE 182
 DB 382 W--VFTGADPRAMGEGTTFVHNS-----PNAMAPLAYDAKLIV--VYPTGV----- 427
 OY 183 TQKGTGRLTLYG-----TNRFAVRPDGTGEIYRHOTLRPRMDDECEFFEMVNTN 232
 DB 428 ---GTP---DIWGDRTLEKERYANSMLAINASTGKLANNPQTHHDLMDMDVPSQSLAD 482
 OY 233 VDVQPSRTEMGLQSLNP--NAATGERVLTGVP-----CKTG--TMQOPDAET 276
 DB 483 IKNAGQTVRAIYVLTGTGNAFVLDR--NGQPIVPEVEKPPQVVKKGPOTKGEYSKIT 540
 OY 277 GEF-----LMA-----RDTNQNMIESIDENGIVYVEDAIL 308
 DB 541 QPFSDNLAPQDKLTDRKDMGATMLDQLMCRVSFKRLANTDGIYTPPSENGTL----- 592
 OY 309 KELDVEYDVCTPLGGRDWPSAALNPD-----GIYFI----- 341

DB 593 -----VFPGMLGVFENGGMGVNPDROVAVMNPGLPFPVSRLLPADPNRAQKAGACT 644
 OY 342 -----PLNNVCYDMMAVDDEFTSMQVNTNNTKLPKGMKIDGIDAIIDISTGRLW--- 393
 DB 645 EQGVQPMYGVY-----GVEISAFSLPGLKPCPKPAMGVYAGVDIAKTHEVYMKR 694
 OY 394 -----SVERRAANVSPYSTGSGVLENGGT--DRYFPALASQETGELTNQTR 437
 DB 695 IGTIRDSLPLNLFQLPVAKIGVPGGSGISFAGVMVFGATQDNYLFAFVNTGKKLMEAR 754
 OY 438 LATVASGQA--ISYEDGMOYVAI--AGGVSYGS 468
 DB 755 LP--AGGQATPMTEYINGKQYVIMAGHSFGR 786

RESULT 13
 DHG_GLUOX STANDARD; PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 CX NCBI_TaxID:442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 of quinoxaline glucose dehydrogenase in Gluconobacter oxydans.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X62710; CAA44594.1; ALU_SEQ.
 CC PIR: S17716; QPKEX.
 CC InterPro: IPR001479; Bac_POO.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
 CC Oxioreductase; POO; Transmembrane; Periplasmic; Signal.
 CC SIGNAL 1 33 POTENTIAL.
 CC CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 CC QUINONE].
 CC FT TRANSMEM 35 54 POTENTIAL.
 CC FT TRANSMEM 59 76 POTENTIAL.
 CC FT TRANSMEM 94 108 POTENTIAL.
 CC FT TRANSMEM 123 138 POTENTIAL.

FT ACT_SITE 470 470 BASE (POTENTIAL)
 PT VARIANT 788 788 H -> N (IN P2 FORM)
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match 9.98; Score 251.5; DB 1; Length 808;
 Best Local Similarity 18.88; Pred. No. 1.6e-11;
 Matches 132; Conservative 77; Mismatches 212; Indels 281; Gaps 24;

```

QY 13 ALALIAAPFAQVPTVDLDEL-----NPPAGWISYQNGENRHSPLTQITTT 61
DB 128 AVLLAFSLFDPDHDSIGELPTQIANASPADPNVASEWMAWYGRGTQAGDRMSPLNQINA 187
QY 62 ENVOGQIDYV-----ANGMGRKV--QVPLIHGVWYLANI----- 96
DB 188 TWSNLKVAWHHTKDMNSNDPGEQTEATPIEFNNLTLYCSLQKLFPAVDGATGVNKKV 247
QY 97 -----VANGVIV----- 103
DB 248 VYDPKLIQINPQHILCRGVSEHETPANAMDSGNPAFTDCAKDSILPVDNGLRVEVDAD 307
QY 104 AGSTC-----QYSPGCFV----- 117
DB 308 TGTCTGCGFNGNCEIDLRVPNPQYTTTGGYEPTSPVITDKLIANSATIDNSVKAQSGA 367
QY 118 -SGHDSATGEELM-----RNYFLPAGSEGETWNGNDYEARMWTGAMGOITVDPVTLVHY 172
DB 368 TQAFVYVYGRKRWVVDASNPDPNOLPDESHVPHSPNSMIVS-----SYDANLNLVYI 422
QY 173 GSTAAGPASEFQKGP--GGTLVGTNTRF--AVRPTGEIVRHOGLTRDMDDEC 224
DB 423 PMGV-----GTPMOWGSDRTKQSERAPGIVALNDTGLKLANFYTTVHHDLMDL 473
QY 225 TFEKAVTNVDVQSTEMEGLOSLNPANAGSERVILGVYCKTGTMQDAETG--EFLMAR 283
DB 474 PSQSLVDVYQKQDGLVPAIYA-----PTKGDIFVLDRTGKEIVAP 517
QY 284 DT-----NYQNMESIDENGIVYVEDALILKELD-----EYD 316
DB 518 ETTPQGAAPADHTSPQPSQ--LTLRPKNPLNDSIDWGTIPDQFHCITYHTLRYE 574
QY 317 -----VCPFLGGRDMPASALNPDGSIYFIFLNNVCYDMAVDO----- 355
DB 575 GPETPSLKGSLIFPGDLMFEMGLAVDPQROVAFANPISLPEVSOLVVRGPNPLMPE 634
QY 356 -----EFTSMQVYNTSNVTKLPPGKDMT-----GRIDAIDISGRIL 392
DB 635 ENAKGTGETGLQHNNGIPIYAVNLHPFLDPVLPFGIKMPCRTPPKGYVAGIDLTKNKV 694
QY 393 WSYERAANYS-----PV-----LSTGGGVLP--NGGTDYFRALSOETG 430
DB 695 WQHRNGTLRDSWYSSLPILPIPIKIGVPSLGGPLSTAGMLGLTASMDYIRAYVLTG 754
QY 431 ETLMOGRLATYVAGSALISYEVDGQVYALAGGVSYSGSLNS 472
DB 755 KYLMODRLPAGQAQATPTTYAINGKOYI-----VYTAGGHNS 790

```

RESULT 14

QY 14 QYIA_AQICA STANDARD; PRT; 809 AA.
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN QYIA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter
 OX NCBI_TaxID=471;
 RN [1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsemore D.A., Ornstun L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quiaA, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RN J. Bacteriol. 176:7659-7666(1994).
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=9601389; PubMed=7592351;
 RA Elsemore D.A., Ornstun L.N.;
 RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus.";
 RN J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
 CC dehydroquininate + reduced pyrroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: PO4.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY. FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PYRROLOQUINOLINE-QUINONE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PYRROLOQUINOLINE-QUINONE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: L05770; AAC37161.1; -;
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 SQ SEQUENCE 809 AA; 88196 MW; 71F67CEBA62BFCB CRC64;

Query Match 8.38; Score 209.5; DB 1; Length 809;
 Best Local Similarity 23.08; Pred. No. 2.6e-08;
 Matches 127; Conservative 59; Mismatches 222; Indels 145; Gaps 22;

```

QY 17 LAAPFAQVPTVDLDELANPAGWISYQNGENRHSPLTQITTENGOLOLV-----WA 72
DB 278 LVQPTLAGATPVA-ALAN-----TECPRYVTNTVDRLIAVANADGCA 320
QY 73 RGMQGRKVQVPLIHGV-----MYLANIVANGVIVAGSTGYSPF-----GCEYS 118
DB 321 RCKDPG-VNGTVNLHEHIGENTAPREYVSAPTIGTITVIGSRADVADMDGCVIR 379
QY 119 GHDSATGEELMRYNFIYPAGEEGDETGNDYEARMWTGAMGOITVDPVTLVH--YGSTA 176
DB 380 AYDVTGRLMA--FDRPNRDPYVYAKPGEIYKRSTNSWAMSYDPOMNTVFLPMGSSS 437
QY 177 VGPASEFQKGPITGLNTRPAVRPDGTGEIVRHOGLTRDMDDECFTFEMVNTNDVQ 236
DB 438 VDWVGNR--TAADHKYNTSV-LALDATGTGKEKVVYNTVHNDLMD-----FDLPQ 485

```

Fri May 24 11:27:37 2002

wallick-934-95.pcp.rsp

Page 13

237 PSYMEGLIOLINPNAATGRRVLTGVPCKTGMMPDAETGEEL-----NA 282
486 PS-----LVDFPKDDTTPKPAVIGT--KSGOFYLDKRYCKPLTKVLEPIKADIPGE 538
283 RDNNYONMIESIDENGIVTNE-----DAILKEID--VEYD-----VC 318
539 QYSTQPRVEMQIGNOLKESDMGATPFQOLMCRINFSKMRDGLYTAAGTDSLSF 598
319 PTFGLGHDMESALNPDGSIYF-----IPLN--NVCYDMMAVD 354
599 PGLSGNMWGSIAFDPTHRMYFVNDRLGLMQLKQTPEDKIQANGGEKYNMGAVP 658
355 QEFTSMQVYNTSVTKL--PPGKDMIGRIDAIDISTGRTLSVEERAAANS----- 403
659 MKGTPYKNNKRRFMSALGIPCOQKPPFGTMAIDMKTRQVAMQVPLGTIODTGPAGIKMGL 718
404 -----PVISTGGGVLFNGCTDRYFRALSOETGETLQTRLATVAGSQAISY--E 450
719 KAPIGMPTIGPMTQGLVFAATQDYLRALNNSNGKELMKARLPVGSQGTSPMSYSP 778
451 VDMQYVAIAGG 463
779 KTKQYVYVVSAGG 791
Db
QY 451 VDMQYVAIAGG 463
Db 779 KTKQYVYVVSAGG 791
RESULT 15
OUIA_XANCIJ STANDARD: PRT: 790 AA.
ID 09XD78;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable quinolate dehydrogenase [pyrroloquinoline-quinone]
(EC 1.1.99.25).
GN QUMA.
OS Xanthomonas campestris (pv. juglandis).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
CC Xanthomonas.
OK NCBI_TaxID=44291;
RN SEQUENCE FROM N.A.
RC STRAIN=CS;
RX MEDLINE=20063481; PubMed=10594704;
RA Lee Y.-A., Lo Y.-C., Yu P.-P.,
RT "A gene involved in quinolate metabolism is specific to one DNA homology
group of Xanthomonas campestris.";
RL J. Appl. Microbiol. 87:649-658(1999).
CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
dehydroquininate + reduced pyrroloquinoline-quinone.
CC -1- COFACTOR: POQ (BY SIMILARITY).
CC -1- PATHWAY: QUINATE ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
PROTOKATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPIATE
PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstations on
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL: AF109471; AAD38453.1;
CC InterPro: IPR001479; Bac_POQ.
CC InterPro: IPR002372; Bac_POQ_repeat.
CC Pfam: PF01011; Bacterial_POQ_5.
CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
CC Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
CC TRANSMEM 22 42 POTENTIAL.

FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 77 94 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT SEQUENCE 790 AA; 82896 MW; B75F29B52A49EE6F CRC64;
SO

Query Match 7.5%; Score 189.5; DB 1; Length 790;
Best Local Similarity 21.2%; Pred. No. 8.6e-07;
Matches 128; Conservative 57; Mismatches 204; Indels 215; Gaps 27;

4 TSLMASAGAL-----ALLAAPFAQVTPVT-----DELLANPAGEWIS--- 43
238 TSSVQRCRGIGYEDADALPAPSVANPSPIAAVYAQAGNCRRRLFTITIGRLIAYDA 297
44 -----YGNQENYRHSPLQITTEVNSQLDLVWARGQPKQVOVTPLIHDGVYLAN 95
298 DTGAFCCGGGNS-----CGVDKAGLGAAP-----DPTQLIS 330
96 --IVANGVIVAGST-----COYSPFGCFVSGHDSATGEELRNRYFIPRAGEEDETWGD 148
331 PPLVAGTVVGGRTADDNVQTDMPGVVVG-----SMWSP--VRSAGLDP-----GNP 377
149 YE-----ARWMTGAMQIITYDPYTNLVHGSTAVGPASETORGTPGTLTYGN-- 196
378 HDRQAPAGSSIVKSTPPNVMAMPMSYDAAMNTVF-----LPLGGPSTDLYGERT 426
197 -----TRFAVRPPTGELIWRHOTLPRDNDQCTFEMMTVNVYVOPSTEMEGLIOSIN 248
427 ALDHRYGASVIALDATTTGAKEKVVQTVHNDLMD-----FDLPMPQSL-----ID 470
249 PNAATGERRVLTGVPCKTGMMPDAETGEELW-ARDT-----NY---QNMIESI 294
471 FPNDDGSHTPAVVIGTGAQIYVLDKRTGKPLTEVREVPKGSIDIAHEQYAPTOPLSYGM 530
295 DENGIVTNEDEILKEIDVEYDVC-----PTFLGDMPSA 330
531 PGIQKHLTESDMGATAMQMLCRIFAQKMRBEGLYTAPGTDVSLSPFGSGGNMGGL 590
331 ALNPDGSIYF-----IPLNN-----VCYDMMAVDOEFTSMQVNTSN 367
591 STDPYHDVVFANDMLGLVMQIPADTRKAEAGGGAIVTGGAVPLKGTPIYAVANKNF 650
368 VTKL-----PPGKDMIGRIDAIDISTGRTLSVEERAAAN-----YSPV----- 405
651 LSAIGIPCOAPP-----YGTLSHIDIKTSIMQVYPVGTQDTGFGIKMHLPIPIGMPTL 706
406 ---LSTGGGVLFNGCT-DYFRALSOETGETLQTRLATVAGSQAISY--EVDGMQYVAI 459
707 GGTLSGGGLVITAGQDYLRALNNSNGKELMKARLPVGSQGTSPITYVSHKTKQYVVI 766
QY 460 AGG 463
Db 767 SAGG 770

Search completed: May 24, 2002, 10:30:50
Job time: 628 sec

Fri May 24 11:27:37 2002

walick-934-95.pep.rsp

Page 14

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:27 ; Search time 93.8 seconds
(Without alignments)
870.508 Million cell updates/sec

Title: WALICK-934-95.PEP

Perfect score: 2530
Sequence: 1 MKPTSLMASAGALLAAP.....GMQYVALGGVSYSGSLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:protist:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*
15: SP:virus:*
16: SP:bacteriophage:*
17: SP:archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	54.5	608	2	093RE9
2	446.5	17.6	698	2	09KH03
3	435.5	17.2	695	2	09PF02
4	430.5	17.0	629	2	09A048
5	428	16.9	601	2	09BYW8
6	427	16.9	601	16	092WY9
7	422.5	16.7	708	2	046444
8	421	16.6	742	2	053362
9	420.5	16.6	601	2	P71509
10	399.5	15.8	691	2	09AF95
11	394.5	15.6	633	2	024759
12	386	15.3	599	2	091935
13	371	14.7	623	2	09AGW3
14	364.5	14.4	695	2	093460
15	359.5	14.2	573	2	059540
16	252.5	10.0	790	2	09X255

17	247	9.8	785	16	098KF6	098KF6 rhizobium 1
18	237	9.4	803	16	091115	091115 pseudomonas
19	236.5	9.3	786	2	P95466	P95466 pantoea cit
20	217	8.6	777	16	092RB3	092RB3 rhizobium m
21	208.5	8.2	182	2	032699	032699 hyphomicrob
22	204.5	8.1	470	2	030326	030326 acetobacter
23	203.5	8.0	179	2	032697	032697 hyphomicrob
24	200.5	7.9	180	2	032696	032696 hyphomicrob
25	200.5	7.9	180	2	032700	032700 hyphomicrob
26	200.5	7.9	181	2	032615	032615 hyphomicrob
27	199.5	7.9	181	2	032692	032692 hyphomicrob
28	197.5	7.8	181	2	032703	032703 hyphomicrob
29	196.5	7.8	181	2	032621	032621 hyphomicrob
30	195.5	7.7	182	2	032706	032706 hyphomicrob
31	193.5	7.6	180	2	032701	032701 hyphomicrob
32	191.5	7.6	182	2	032612	032612 hyphomicrob
33	191.5	7.6	184	2	032693	032693 hyphomicrob
34	190.5	7.5	172	2	033894	033894 hyphomicrob
35	187	7.4	639	2	P77931	P77931 pseudomonas
36	186.5	7.4	172	2	033882	033882 methylosinu
37	186.5	7.4	179	2	032707	032707 hyphomicrob
38	185.5	7.3	180	2	032694	032694 hyphomicrob
39	185.5	7.3	180	2	032704	032704 hyphomicrob
40	185	7.3	171	2	033884	033884 methylosinu
41	183.5	7.3	179	2	032702	032702 hyphomicrob
42	182.5	7.2	184	2	032695	032695 hyphomicrob
43	181.5	7.2	185	2	0918K6	0918K6 uncultured
44	180.5	7.1	644	2	052551	052551 pseudomonas
45	176.5	7.0	172	2	093K58	093K58 methanotrop

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT: 608 AA.
AC 093RE9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucosobacter saccharoferrogens.
OC Bacteria; Pseudoglucosobacter.
NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IFO 14464;
RA Shidara T., Saito Y.;
RT "Alcohol dehydrogenase."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1; -
SQ SEQUENCE 608 AA; 65101 MW; 0AC9C97AE11BA570 CRC64;

Query Match	54.5%;	Score 1380;	DB 2;	Length 608;
Best Local Similarity	46.7%;	Pred. No. 1.2e-93;		
Matches 272;	Conservative	58;	Mismatches 132;	Indels 120; Gaps 5;
OY	10	SACALALLAAPFAQ-----VFPTDELANPAGEWISYQNDEN	50	
OY	1			
DB	16	STALLIASLGPFAFDHANAAPESKAGQSAIENFQVPTADDLAKNPWPTLRGNYOG	75	
OY	51	YRHSPLTOTTETWVGOLVWARGMPGKVOVTPPLHIDGVNYLAN-----	95	
OY	76	WGSPPLDQINKDVGDLOLWMSRTMEPSNDEGAIAVNGVIFLANTNDVQALDKTGS	135	
DB	76	WGSPPLDQINKDVGDLOLWMSRTMEPSNDEGAIAVNGVIFLANTNDVQALDKTGS	135	
OY	96	-----	95	
DB	136	IMFYRRRLPSASFINSLSGAARSTALRGDKYFVSWDNFVALDARTGLANETNRGCG	195	
OY	96	-----IVANGVYAGSTQYSPGCVSGSDSATGEELMRNYFIPRAGEEGDET	144	

```

Db 196 VEEGVANSSGPIVVDGVVAGSTCGSGGECVTGTGDASSGELMRTFIPRPEGGDT 255
QY 145 WGN-DYEARMMTGMCOITDPTNLVHYGSTAVGASSETGPGTGGTLYGTNRFAVR 203
Db 256 WGNAPYERMMTGMCOITDPTNLVHYGSTAVGASSETGPGTGGTLYGTNRFAVR 203
QY 204 DTGEIVRHOTLPDNDDECTEFEMVNTVDVOPSTMEGLQSTNPNAAATG-RRVLTGY 262
Db 316 KTGEEVWKHOTLPDNDDECTEFEMVNTVDVOPSTMEGLQSTNPNAAATG-RRVLTGY 262
QY 263 PCKTGTMOFADTGEELFARDTNQNMIESTIDENGITVVEDALIKELDEVDVOC 322
Db 376 PCKTGTMOFADTGEELFARDTNQNMIESTIDENGITVVEDALIKELDEVDVOC 322
QY 323 GGDWMSALNDPDSGIYFIPLNVCYDMAVDOFTSMQVNTSVTKLPCKGDMIGRID 382
Db 436 GGDWMSAGYILKPSMLYVPLSMCYDMATTEATPADVNTDAILVLPCKTNGRVD 495
QY 383 AIDISTGRTLMSEVERAANYSPVLTSGGVLENGTDYFRALSOFTGETLMOPLATA 442
Db 496 AIDLATGETKMSYETRALDLPVLTGGDLVFGGIDRFRALDSESGEWMSTRLPNAV 555
QY 443 SGOAISYEVDMQYVAIAGG-----VSYGSLN 471
Db 556 SCTTTSYIDGRQYVAIVVSGSLGPTFGPTFPDVDSAGAN 597

```

RESULT 2

```

Q9KH03 PRELIMINARY; PRT; 698 AA.
ID 09KH03;
AC 09KH03;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE TERAHIDROFUFURYL ALCOHOL DEHYDROGENASE.
OS Alcatigenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21122557; PubMed=11222593;
RA Zarnit G., Schrader T., Andreesen J.R.;
RT "Catalytic and Molecular Properties of the Quinohemoprotein
RT Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo.";
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL: AF277373; AAF86335.1; -.
DR HSSP: 092477; 1FLG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_PQQ; 6.
DR Pfam: PF00034; Cytochrome_C; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F18DBD12FB2 CRC64;

```

Query Match 17.6%; Score 446.5; DB 2; Length 698;
 Best Local Similarity 25.5%; Pred. No. 1.3e-24;
 Matches 149; Conservative 74; Mismatches 203; Indels 159; Gaps 21;

```

QY 10 SACALALILAPAFV-QVTPTDELNPAG--EWISYQONENYRHSPLTOITTEVNG 65
Db 14 AAASVALPAPAFGANNAARVDGAIRANEGATPMPNSYGLDAEYERFKELEQVNAANVR 73
QY 66 QQLVWAKMPGK-VQVTPLIHDGVNVL----- 93
Db 74 NGLIAMSYLESTRGVEATPLVVDGVVYVSAFMSVVAHALDARTGRLTWYDPOVRODAY 133
QY 94 -----ANIVA 98

```

```

Db 134 KCCDDVNVNRGVALYQKVFVGAFDGLVAIDAATGKRVMEODTLVDRSKSYTTIGAPRY 193
QY 99 NG-VIVAGSTCOYSPGCGVSGHDSATGEELMRYFTIP-----RAGEEDETWGD 148
Db 194 NGKVIINGSGAEVGVNG-VITAYDAETGSKQOMRYTLPDGPAPRFENMAKAAATW--D 250
QY 149 YEAR-WMTG----ANGQITYPEVNTLVHYGSTAVGASSETGPGTGGTLYGTNRFAVR 202
Db 251 PSCKYWINCGGCTVNTAFNPDELMTIGTGNAGPSKRLRSPKGGDLVQAAAY-VALN 309
QY 203 PDTEIVRHOTLPDNDDECTEFEMVNTN-VDOVOPSTMEGLQSTNPNAAATGRRRLT 260
Db 310 PDTEIVRHOTLPDNDDECTEFEMVNTN-VDOVOPSTMEGLQSTNPNAAATGRRRLT 260
QY 261 GVPCKTGTMOFADTGEELFARDTNQNMIESTIDENG-ITYVEDALIKELDEVDVOC 318
Db 352 NAF-KNGEFFVIDRTNCKFISAKNFPVDMASGYDKNGRPETPDADTSGKPA----DVY 406
QY 319 PTFEGGRMPAALNDPDSGIYFIPLNVCYDMAVDOFTSMQVNTSVTKLPCKGDMIGRID 382
Db 407 PGPFGAHNMHSMFHPRLGLAFIPQHYPL-TLADNKEM-----VHNKDSPEAHRYGN 461
QY 368 -----VTKLPCKGDMIGRIDALIDISTGRTLMSEVERAANYSPVLTSGGVLENGTDYFR 423
Db 462 LGMLVNAEPSPRSKPMGRLLAMPDLQAKVNRHDAHGPWNGSTLATAGNLVFOGTADGRLY 521
QY 424 ALSQETGETLMOPLATVASGAISYEVDMQYVAIAGGVSYG 467
Db 522 AYHAATGETKMOAPFGSGVVAAPVYTLIDGRQYVAVVSGSLGPTFGPTFPDVDSAGAN 597

```

RESULT 3

```

Q9F902 PRELIMINARY; PRT; 695 AA.
ID Q9F902;
AC Q9F902;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudomonas stutzeri (Pseudomonas perfectionarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-2;
RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainerd J.B.;
RT "Identification and characterization of genes activated by 2-
RT chloroethanol in Pseudomonas stutzeri BC-2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF176640; AAG09249.1; -.
DR HSSP: 092477; 1FLG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_PQQ; 6.
DR Pfam: PF00034; Cytochrome_C; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

```

Query Match 17.2%; Score 435.5; DB 2; Length 695;
 Best Local Similarity 25.0%; Pred. No. 8.2e-24;
 Matches 140; Conservative 72; Mismatches 207; Indels 141; Gaps 18;

```

QY 15 ALAPAFVQVTPDEL--ANPAGEWISYQONENYRHSPLTOITTEVNGQLOLWNA 72
Db 16 ALIVA-AGAQAANVDAIRASDQSEWLSHGRTVAEORFSPKQIDAGNVKTLANAY 74
QY 73 RGNPCK-VQVTPLIHDGVNVL----- 94

```

Db 75 LDLENNRGLEAPLVSDVLYASLSWSRVMAVDRSGKRLMOEDPDVDRGSHSKYCCDAV 134
 QY 95 -----NIYANGTVAG 105
 Db 135 NRGVALNNGRVYGAIDGRLIALDAKTGRELMSBQCTDPKPYISITGAPRVYAKVLIEN 194
 QY 106 STCOYSPFGCVGSHSATGEELMRNYFIPRAGE-----EGDETV-GNDYEARM 153
 Db 195 GGAETGVYRG-FESAYDAETGKMAKRYVTP-GDPAQYEHPELAEAKTKWGDQY---W 248
 QY 154 MTG-----ANGCITPDVNTLVHGSTAVGPASSETRGTPGCTLTGTNRFAVRPDGEIV 209
 Db 249 KLGGGGVWDSMAIDPELDLITGTGNGSPWNRKISPPGGDNLYLSITLARDSSKLL 308
 QY 210 WRHQTLPDNDODCTEFEMAVTVNVOPTSTEMGLOSINPNAATGERVLTGVDPCKTGM 269
 Db 309 WHYQTPGTEWDFATODITLATL-----ELDG-----KPRKLMQAP-KNGFF 351
 QY 270 WQDPAETGELMARNTYONMIESID--ENGIVYVNEALIKELDYEXDVCTFLGGRW 327
 Db 352 YVLDRTGELLASKEFKVTAKEVLDATGRPEVPGSRYEKQVVM---PSSFGAHNW 408
 QY 328 PSALNPDGSIYFPLNVCYDMAVDOEFTSMQVNT---SNVTKLPBGKDM-GRID 382
 Db 409 HSMSPNQIGLWIPYQELPGVYRNEGATFKKIDGLNTGTGFSQTHIIP--RDVAGSALL 466
 QY 383 AIDISTGRITMSVERAANYSPVLTSGGVLFPNGCTDRTYRAISOEGETLMOTRLATVA 442
 Db 467 AMDVYRREMARVPHSFYNNGGTLTAGNLVFOGTADQULHAYSADGQALNMFQAOTGI 526
 QY 443 SGQAISTEVDGQYVAIAGG 462
 Db 527 VAARISFSLDGEQYVAVMAG 546

RESULT 4
 ID 09A048 PRELIMINARY; PRT; 629 AA.
 AC 09A048;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE METHANOL DEHYDROGENASE MAA'.
 GN MAA'.
 OS Methylobacterium nodulans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OC NCBI_TaxID=114616;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-ORS2060;
 RC MEDLINE-2056686; PubMed-11114919;
 RA SY A., Girard E., Jourdan P., Garcia N., Willems A., de LaJudie P.,
 Rabin Y., Neyra M., Gillis M., Boyvin-Wasson C., Dreyfus B.;
 RA "Methylobacterium nodulans" bacteria nodulate and fix nitrogen in
 RT symbiosis with legumes";
 RT J. Bacteriol. 183:214-220(2001).
 RL EMBL; AF220764; AAC49450.1; --
 DR HSSP; P38539; 4AAH.
 DR InterPro; IPR001479; Bac_P00.repeat.
 DR Pfam; PF01011; Bacterial_P00_7; 1.
 DR PROSITE; PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
 DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
 SO SEQUENCE 629 AA; 68798 MW; 3DB559D8373BB0B2 CRC64;

Query Match 17.0%; Score 430.5; DB 2; Length 629;
 Best Local Similarity 24.0%; Pred. No. 1,7e-23;
 Matches 143; Conservative 88; Mismatches 198; Indels 167; Gaps 21;
 4 TSLIMSAGALIALAPAFQVTPVTDLL-ANPPAGEWISYQONQENYHSPLTQITTE 62

Db 11 TGVSYAALALLPLAGS-----ALANDKIVLELSKSDGNWVWPKNYSDNTSKLKQINAE 66
 QY 63 NGGOLQVW--ARGWQPGKVQVTPILHDGYMTL----- 93
 Db 67 NYKNLKVSMQFSTGLNGH-EGAPLVVDGIMVYHTSPNNTFALSGDDPKILMDDKPKQ 125
 QY 94 -----ANIVA-----NGVIVA 104
 Db 126 NPAARSVACDVLNNGLAIVWPDGKTPSLIKTLIDGHVVALNAQGTETWIKIENSDIRV 185
 QY 105 GSCQXSPF-----GC-----FVSGHDSATGEELMRNYFIPRAGEGD----- 142
 Db 186 GSTITLAPVYVKQVIGSSGAELEGVRLTAYDVRIGCEQKWRV--ATGPPSDILLAK 242
 QY 143 -----ETWGNDEARWMTGA--WGQITPDVNTLVHGSTAVGPASER 183
 Db 243 DENIHNAHYGKGLGISTWEGD--AMKIGGGTNGWAVADPGNLTLYFGTGNPAPNNE 299
 QY 184 ORGTPGGTLTGNTREAVRPDGEIYWRHQTLPDNDODCTEFEMAVTVNVOPTSTEMEG 243
 Db 300 MR--PGDKW-TMTIFARDVDTGEARFGIOKTPHDEMVDAGVNVML-----STOKD- 348
 QY 244 LOSINPNAATGERVLTGVDPCKTGMQDPAETGEFLAR---DTYONMIESIDENGIV 300
 Db 349 -----RSGKERKLLHPDRNGIYVTLDTGTGDLISAKIIDT--VAVRTYDLKSGL 398
 QY 301 TVNEALIKELD-VHYDVCPFLFGGRWPSALNPDGSIYFPLNVCYDMAVDOEFTS 359
 Db 399 PVADPEYGTMDHLKNDICPSAMQHNGHSDYDEPKLEFYMGINICDMDEPFLPRA 458
 QY 360 MDVYNTSVNTKLPCK-----DMIGRIDIDISTGRITMSVERAANYSPVLTSGGVL 413
 Db 459 GQFVGTATLNMVDPKGRONAEGLQIKAYDALTGTFKKEKMERFAPVWGCTLATAGNV 518
 QY 414 FNGGTRIFPALSQETGETLMOTRLATVAASQAISTEVDGQYVAIAGGVSQSG 469
 Db 519 FYGLTDGFIKARHSDIGELIMKAKLPSGALGYVYTHKGTQYVAI-----YTGVG 569

RESULT 5
 ID 09EYW8 PRELIMINARY; PRT; 601 AA.
 AC 09EYW8;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
 GN MAA'.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OC NCBI_TaxID=382;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BM1021;
 RC Fennel B.J., Tiwari R.P., Dilworth M.J.;
 RA "Regulation of C1 assimilation in Sinorhizobium meliloti";
 RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 RL EMBL; AF309488; AAC31643.1; --
 DR HSSP; P38539; 4AAH.
 DR InterPro; IPR002372; Bac_P00.repeat.
 DR Pfam; PF01011; Bacterial_P00_7; 1.
 SO SEQUENCE 601 AA; 65686 MW; 09F74372B2E2A83 CRC64;

Query Match 16.9%; Score 428; DB 2; Length 601;
 Best Local Similarity 24.9%; Pred. No. 2.4e-23;
 Matches 149; Conservative 72; Mismatches 202; Indels 176; Gaps 22;
 14 LALLAAPAFQVTPVTDLL-ANPPAGEWISYQONQENYHSPLTQITTEVNGQLOLVW 71
 Db 8 LAIMISGGQAVAFANDELQKLDLP-NQMAIQTGYANLRYSKLDQINKDNVGLQVAVW 66

Db 8 LAIMISGGGAOYAFANDELQKLIIDP--NOMAIGTGYALNRKSLDOIINDNKGKIOAV 66
 QY 72 A-----RMO-----PGVVQVPLIHH----- 88
 Db 67 TESTGVLRGHEGSSPLVIDLMTYHTPPRNTYVALDLSKGGIWKYERKODPVNIPVCC 12
 QY 89 -----GVMLN-----IVANGVIGST----- 10
 Db 127 DTVNGVAVYADNKKFLHQADTVVALDAATGKVIWVKNGDATGENTATVMPYKDKL 18
 QY 108 --CQSPFGC--FVSGHDSATGGEELMRNFI-----PAGEBDETW 14
 Db 187 VQISGGEVGVGHVTAASMADEKVLIRGYSKMPDSDLLIDPEKTHIGKRPVKRDSGLTWT 24
 QY 146 GNDYFARMWTG---AMQIITVPYVNLVHSGSTAVGASGTQNGTGGTLVYNTNFAVR 20
 Db 247 EGD---OMKIGGGTWTGWVSYDEENLVYTGTPSTWNPJOR--PGDNRW-SMTIFARD 30
 QY 203 PDTEGIYWRHQTLPRNDMDECTFFPMATYNDVQSTEMEGLOSIIPNMAAGERRVLGV 26
 Db 301 VDTGMAKWLYQMTPHDEWYDQVEMEMILTEQIDGK-----DKLLIHF 34
 QY 263 PCKTGIMQPDFATGEBFT-----NARDINYQNMIE-----SIDENGIYV 30
 Db 345 D-RNGGYTMDRYTEGELLVAKRYDPVTYNNATEYVMDPKSDKGRPOVVAQYSTEQN--- 40
 QY 302 VNEDAILKELDEVYVCTPFGGGDMFSALANDSGITFIFPLNNVCYDMAVQDEFTSMD 36
 Db 401 --EDTNTT-----GVCPALAGTKRDOQPAVYSKTELEFVPLTNVHCMDYDPFNVSTAAQ 45
 QY 362 VYVTSNVTKLPPGKDM---IGRIDAIDISTGRTLSVERAANYSPVLTSGGVLPNGST 41
 Db 453 PTVGATLMIYRP-KDSHGKGNFIANDNKEGIKMSLPEPFSVMSGALATAGVYFTGL 51
 QY 419 DVEFALSLQELGELFMOTRLATIVASGQASITSEVDGMQVVAIAG-----GVSYSGGLNS 47
 Db 512 EGYLAAVDAATGAKELRYKFKTPSGVIGNVMTYARBGKQYAVVLSGCGMAGIGLAAGLTN 57
 RESULT 7
 Q46444 PRELIMINARY: PRI: 708 AA.
 AC Q46444;
 DT 01-NOV-1998 (TREMblrel, 08, Created)
 DT 01-NOV-1998 (TREMblrel, 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel, 19, Last annotation update)
 DE OUNIOHAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 (EC 1.1.99.-) (OH-EDH1).
 GN OS
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
 OX NCBI_TaxID:285;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15667;
 RC MEDLINE=96184549; PubMed=8654419;
 RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinoxinohemoprotein ethanol
 dehydrogenase of Comamonas testosteroni";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RP SEQUENCE OF 32-54 AND 477-490.
 RC STRAIN-ATCC 15667;
 RX MEDLINE=95324580; PubMed=7601151;
 RA De Jong G.A.H., Geerlof A., Stoorvogel J., Jongejan J.A., De Vries S.,
 Duine J.A.;
 RT "Quinoxinohemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 RT Purification, characterization, and reconstitution of the apoenzyme
 with pyridoxalminol quinoxin analogues.";
 RL Eur. J. Biochem. 230:899-905(1996).
 RP [3]
 RP CHARACTERIZATION.

RA MEDLINE-86242113; PubMed-3521592;
 RX Groen B.W., van Kleef M.A., Duine J.A.;
 RT "Quinolhaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 testosteroni".
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION.
 RA MEDLINE-21356088; PubMed-11679760;
 RX Oubrie A., Hultinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 RA Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinohaemoprotein alcohol dehydrogenase from
 Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: PQQ, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 CC EMBL: X81880; CA57464.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt.C1.
 DR InterPro: IPR002329; Cyt.C1C.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR Pfam: PF00034; Cytochrome_c_1.
 DR PRINTS: PR00605; CYTOCHROME.C1C.
 DR SIGNAL: PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 KW SIGNAL: 1 31
 FT CHAIN: 32 708 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
 FT TYPE 1.
 FT BINDING: 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING: 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL: 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT SEQUENCE: 708 AA; 76822 MW; 99AB54BDD6ACCA3 CRC64;

Query Match 16.7%; Score 422.5; DB 2; Length 708;
 Best Local Similarity 23.5%; Pred. No. 7.7e-23;
 Matches 139; Conservative 79; Mismatches 222; Indels 151; Gaps 18;
 3 PLSLMSAGALLAALPAFAQVTPVT-----DELANPP-AGEWISYGOQEN 50
 12 PERMWLLAACIG--SAAAFQGTGPAQAQAAAVQVDGFIKANAARPTIGVYAE 69
 51 YHSPLEQTITENYGOLOLVARQMOPK-VQVPLIHGVMLA----- 94
 70 TRSRDLQINANAYKDLGSLANSYNLESTRGVEATPVVDGIMVYSASVVAIDRTGN 129
 95 -----NIVANG----- 100
 130 RIWYDPOIDRSTGFGCCDVRGVALMKQYVVGAMQRLIALDAATGKEVWHQNTFE 189
 101 -----VIVAGSTCOYSPGCFVSGHDSATGEELMRNYFLP----- 135
 190 GQKSLITIGAPRVKGVYIKKRGAEYVGR-YITAYDAETGERKWRMSVPGDPSKRF 248
 136 --RAGEGDETDGNDYEAMWTGA-----WGQITVDPTNLVHYSIAVGPASSETORTP 188
 249 EDESMKRAARTW--DSGKWMEEAGGGGTWDSMTFDALNMYGTGSGPSMKYRSPK 306
 189 GGTILGNTFAVRPDTEIWRHQTLPNDNDOECTEFEMAVTVVDQPSRMEGLQSI 248
 307 GGNLYIASIALDPDQTKYKMYHETPGDNDYDSTPMLADIKT----- 353
 249 PNAATGERRVLTGVPCKTGMQFDATGEFLIARDIVYONNMIESIDENG-IYTVNEDAI 307
 354 --AKPKRVILHAP-KNGFFVLDRTNGKFKISAKNFVYVNWASGYDRHGKPIGI--AA 406
 308 LKELDEVDCPTFLGGRDWPSSALNDGSIYFIPLNVCYMMAVDQ-EFT----- 358

DB 407 ARDSKPODAVPGPYGAAHNMHSPNFDPQITGLYLLPAQNYPVNLMDKKMEFNQAPGPKQ 466
 359 SMDVYNTSNVTKL-PGCKDMIGRLDAIDISTGRTLSVERAANYSPVLSTGGVLENGC 417
 467 SGTGMNTAKFFNAEPKSPKPRGLAMDPAQAQAAMSVHEVSPMNGGTLTLAGNVYDGT 526
 418 TDRYRALSOETGETIMQTRLATVAGSAISYVDGMOYVLA-GGGSYSYG 467
 527 AGRLVYHAATGKLMKAPGTGVVAAPSTYMWDRQYVSAVNGGYG 577
 RESULT 8
 053362 PRELIMINARY; PRT; 742 AA.
 ID Q53362; Q44159;
 AC 01-NOV-1996 (TREMBLER, 01, Created)
 DT 01-NOV-1996 (TREMBLER, 01, Last sequence update)
 DT 01-OCT-2001 (TREMBLER, 18, Last annotation update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
 OS Acetobacter pasteurianus (Acetobacter turbidans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_Taxid=438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:NC11380;
 RX MEDLINE=94042848; PubMed=8226628;
 RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
 RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
 pasteurianus".
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL: D13893; BAA04252.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR Pfam: PF00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR SEQUENCE: 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 16.6%; Score 421; DB 2; Length 742;
 Best Local Similarity 24.7%; Pred. No. 1.1e-22;
 Matches 150; Conservative 72; Mismatches 205; Indels 180; Gaps 20;
 11 AGALALLAARPAFAQVTPVDEL--ANPPAGEMISYGOQENYHSPLEQTITENYGOLO 68
 23 AALPYAAVAPARADGGGNTGEALIHADHPENWLSGTYSQRYSPLDQINRSNGDLK 82
 69 LVARQMOPKQVQ-VPLIHGVML-----ANI----- 96
 83 LAWYTTLDNRGQATPLVVDGIMVATNMKWEALDAATGRLMWDYKPKYGINADKGC 142
 97 ----- 96
 143 CDTVRKAGYNGKVFWMGTEDGRLVAADAKTGKWEVNTIPADASLKORSYTVDAVR 202
 97 VANGYIVGSTCOYSPGCF--FVSGHDSATGEELMRNYFLPRAAGEGD----- 142
 203 VARGVLVGN--GSEFPAARGVSAFDAETGKLKMRYYTVPNKNKNEPDAVADVNLMSKA 260
 143 -ETWMDYEARMT-----GARGQITVDPTNLVHYSIAVGPASSETORTP 193
 261 YKTWGP--KGAVVROGGGTWDSLVDPVSDLY--LAVG-----NGSPNNKRYRSE 309
 194 --GTN-----TRFAVRPDTEIWRHQTLPNDNDOECTEFEMAVTVVDQPSRMEGLQSI 247
 310 GIGSNLFLGSIYALKRETEYVHFGQATPMDQWYTSVQOIMTLDMFY----- 357
 248 NPNAATGE--RVLTVGPCKTGMQFDATGEFLIARDIVYONNMIESIDE-NGITVNEED 305
 358 -----NGEMRHVILHAP-KNGFFVLDRTNGKFKISAKNFVYVNWASGYDRHGKPIGI--AA 411

```

QY 306 ALKELDEYDVCPFLGDRMPGSAALNPDGTYFPLNNVCY-----DMA 352
Db 412 GLYTLGKFWYIGIPGLGANNHGMAYSPKTHLVYPAHQIFPGKKNVGGEKPPHDMN 471
QY 353 VDOETSDMYNTSNATKLPKPGDMIGRIDAIDISTGRTLMSEVEAANAANSPYSTGGV 412
Db 472 VGLDMTKKGLDPE-ARTAYIKDLHGLMLANDPVKMEIYKIDHKGWNGCVLATGDL 530
QY 413 LFNGGDRYFRALSOETGETLMOTRLATVASQAISTEYDQMOTYA-----IAGG 462
Db 531 LFQGLANGEFHAYDATNGSDLYKFPDQSGIATSPMTYSVNGKQYVAVEVMGCIPISWG 590
QY 463 GVSYGSG 469
Db 591 GVGRTSG 597

```

```

RESULT 9
F71509
ID P71509 PRELIMINARY; PRT: 601 AA.
AC P71509;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE METANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
GN MXAP.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AMI;
RA MEDLINE=97312011; PubMed=9168622;
RA Christoserdova L., Lidstrom M.E.;
RT "Molecular and mutational analysis of a DNA region separating two
RT methyloctrophy gene clusters in Methylobacterium extorquens AM1."
RL Microbiology 143:1729-1736(1997).
DR EMBL: U72662; AAB58890.1;
DR HSSP: P38539; AAAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ.
SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CB8239 CRC64;

```

```

Query Match
Best Local Similarity 16.6%; Score 420.5; DB 2; Length 601;
Matches 147; Conservative 90; Mismatches 203; Indels 157; Gaps 24;

```

```

QY 1 MKPSTILWASAGALALLAAPAFQVTVVDLIANPAGEMISYQONENRHSPLQIT 60
Db 1 KRAVHLALAGGLAA--ASPLANESVLG--VANP--AEGLQTVVDYANTRYSKLDQIN 54
QY 61 TENGOLOLVW--ARGMOPKQVOVPLIDHGYWL--ANIV-----97
Db 55 ASNVKNLQVAMTSTGLRGH--EGSPLVYGNIMYHTPPNIVAYALDLOGAKIWKYEP 113
QY 98 -----ANGVIV-----106
Db 114 KQPSVIVPMCCDTVRKGLAYADGAILLHOADTLVSLDAKSGKVMNSYVANGDPSGETN 173
QY 107 TCQYSP-----FG--CFVSGHDSATGEELMKRYFT-----134
Db 174 TATVLPVKKDVLYIGISGFEVQCHYATAYDLKSGKKWRRYSIGPDOLIVPEKTSIG 233
QY 135 -PRAEGEGETMGNDYERAMKMG--AMGOITVDPVNLVHYGSTAVGASSETQCTPGG 190
Db 234 KPIADSSIKTWEED--QMKTGCGCTMGWFSYDPKIDLMYISG--NFTWNPKQRPQD 288
QY 191 TLYGTNTRAVAPDTEGLVWRHQLPRDNDQCTEAMNTVNDVQPSSTEMGLQSIIPN 250
Db 289 NKM-SMTIMARNPDIGAKAKVYIQMTPHDEWDFGINEMILTD-----QKFDG-----334

```

```

QY 251 AATGERVLTGVCCKTGMMDAETGEFLARDTN-YQNNIESID-ENGIVTNEADLI 308
Db 335 ---KDRPLLHPD-RNGFGTTLDRATGEVLAEKFDPPVNMATVDLDKSKTYGRPLV 390
QY 309 KELDVEID-----VCPFLGDRMPGSAALNPDGTYFPLNNVCYDMAAVDEFTSM 360
Db 391 SKYSTQONEDNVNSKICCPAALGTQDQPAFSPKTELFYPTNHVCMDYEPFRVYTPG 450
QY 361 DVYNTSNVTKLP-PGK-DMIGRIDAIDISTGRTLMSEVEAANAANSPYSTGGVLFNGT 418
Db 451 QPYVATLTMYPAPSGHCKGNFTAMDNLQGIKWSNPEQFSAMGALATSGDVFYGTIL 510
QY 419 DRYFRALSOETGETLMOTRLATVASQAISTEYDQMOTYA-----GVSYGSL 470
Db 511 EGFLEKAVDSKTKGKELKFPKFPSSGILGNVMTYEHKQKHVAVLGSGVAGIGLAAGL 567

```

```

RESULT 10
O9AF95
ID O9AF95 PRELIMINARY; PRT: 691 AA.
AC O9AF95;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE BDH.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Vanaul A.S., Arp D.J., Sayavedra-Soto L.A.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF355796; AAK27220.2;
SQ SEQUENCE 691 AA; 75070 MW; 4FCFDD20CA1AE64 CRC64;

```

```

Query Match
Best Local Similarity 15.8%; Score 399.5; DB 2; Length 691;
Matches 137; Conservative 74; Mismatches 191; Indels 147; Gaps 19;

```

```

QY 39 GEMSTYQONENRHSPLQITTEENVGOLVWARGMOPKQVOVPLIDHGYWL-----93
Db 41 GEMRTHGYDAGTRYSPLAQITPDNAKELGLVMSIDLESSRGVATPIVVDGYVYTAEP 100
QY 94 -----ANIVANGVIV-AGSTCOYSPFGCFVSGHDS 122
Db 101 SYVHALDVRSGKRLMTYDEVRERKGNACCDYVNRGVAHEGKVPFVSLDGRLVA-IDA 159
QY 123 ATGEELW-RNYFT-----PRA--GEGEGETMGNDYERAMKMGAM-----GO 160
Db 160 RTGKRWERNTLIDDKPYITIGAPRYIKGKVIVGNGAEGVGYITAYDPTAASRGV 219
QY 161 I-----TYDPVTNLVHYGSTAVG-----178
Db 220 VREPQDPSLPEDDASMEAAKTMDPAQVILGSGRRRGVGLQYLRKAGFCCTAPAPS 279
QY 179 PASETQRTPGTLYGTNTRAVAPDTEGLVWRHQLPRDNDQCTEAMNTVNDV--Q 236
Db 280 PMSHRKRSAPGDMNLTAISYALRPDTEGLVWRHQLPRDNDQCTEAMNTVNDV--Q 236
QY 237 PSTMEGLQSIIPNAANGERRVLTGVCCKTGMMDAETGEFLARDTNQNNIESIDE 296
Db 340 P-----RKVILHAP-KNGFFVYIDRTDKGKFLSAQNFVPMNATGYDE 380
QY 297 NGIYTVN-EDAILKELDEYDVCPFLGDRMPGSAALNPDGTYFPLNNVCYDMAAVD 355
Db 381 NGRPIENPEGAMGHLSMR--PAPSARTNHSMSYSPTQGLAYPPAQNIPL-VLOEDK 436
QY 356 EFTSMYNTSN-----VTKLPKPGDMIGRIDAIDISTGRTLMSEVERAA 400

```

```

DB 437 NMS-----YNQAGQAGMAGIGNMLGNVPRPBPQSPGRILANDPVOQEVNRKREHVSF 492
OY 401 NYSVLTSTGGVLFNGSTDRYFRALSOETGETLWOTRLATVASSQALSYVDGQVYAA 460
DB 493 WNGGLTVAGNVFQGTADARLLAFAPARDGKELMSAPMGTVIAAPVTEYDQKQVISA 552
OY 461 -GGGVSYGS 468
DB 553 VGMGVGYGN 561

RESULT 11
ID 024759 PRELIMINARY: PRT: 633 AA.
AC 024759;
DT 01-JAN-1998 (TREMBLER, 05, last sequence update)
DT 01-JAN-1998 (TREMBLER, 19, last annotation update)
DT 01-DEC-2001 (TREMBLER, 19, last annotation update)
DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN MXAF.
OS Hypomicrobium methylavorum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.
OC Hypomicrobium group; Hypomicrobium.
OX NCBI_TaxID=84;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM2;
RA MEDLINE=97457202; PubMed=9311140;
RA Tanaka Y., Yoshida T., Matanabe K., Izumi Y., Mitsunaga T.;
RT "Cloning and analysis of methanol oxidation genes in the methylotroph
RT Hypomicrobium methylavorum GM2."
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL: AB004097; BAA3272.1;
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
DR Signal: oxidoreductase.
KW SIGNAL.
FT CHAIN 35 633 POTENTIAL.
SQ SEQUENCE 633 AA; 69653 MW; B47A23A279E2C1B CRC64;

Query Match 15.6%; Score 394.5; DB 2; Length 633;
Best Local Similarity 23.6%; Pred. No. 7.6e-21;
Matches 140; Conservative 88; Mismatches 213; Indels 151; Gaps 21;

OY 6 LLMASGALLALAAPAFQVTPYTDLLANPPAGE-WISYGOENQENYRHSPLTOITTEV 64
DB 12 LMSASCAIIVALQVASSAVANDKILSLSENENWMPGKNYSSTOVNAENV 71
OY 65 GOLULVNA 72
DB 72 KOLKAMSFSTGELHGEAPLIVGDMYVHSSFPKTFALNDPQHILMOHSPKDDPA 131
OY 73 -----RGMO--PGKVQYPLI---HDGVYTLANI-----VANGYIVAGST 107
DB 132 ARSVACDLVNRGLAYWPGDKPTALVITKQLDGLHVALNAKTGEEFVKVENGDIKVGOT 191
OY 108 COVSPGCF-----VSGHDSA-----TGEIMRANFYI----- 134
DB 192 LTPQAPVVDLAIYSGSAGELGVGRHAYANVKTIGEQARKYVATGDEELGLADDFNSAN 251
OY 135 PRAGEE--GDETMGNDYEARMWTCG--WGQITDPTVNLVHYGSTAVGASPASETQRTGTPG 189
DB 252 PHYGQKGLGTATWEGD--AMKIGGTWGMWYAYDPOANLIYSGSNAPAWNMTAR--PG 306
OY 190 GLIYGTTRFAPVPTGELVWRHQLPRDNDQECTEFEMVTVNDVOPSTEMGLQSLINP 249
DB 307 DNKW-TWITIRADADTGKMKFGYKTPHDEMDPAGVNYIMLSE-----QTDKE----- 353

```

```

OY 250 NAATGERVLTGVPCKTGTMQPDATGEELMA---RDTNVMQMIIESIDENGIYTVNEDA 306
DB 354 ----GKKRLILHPDRNGIYVTLDRNGDLISAKLDDT--VNVFKHVDLKSGLPYRDE 407
OY 307 ILKELIVE-YDVCPFLIGRDPWSPALNPDGIFYFIPLNVCYDNMAVQDEFTSMVYNT 365
DB 408 FGRMDKRGTEICPSAMGYHNGCHSDYDPTKQLPFMGIGINIMQMEFPLPYAAGQFVYG 467
OY 366 SNVTKLIP-PGKDM-----IGRIDAIDISTGRTMSEFRAANYSVLTSGGVLFNGND 419
DB 468 ATLMMYGPGRGDRONTLYLGQIKATVATNTYTKMEHMERFSVMGTLATAGNLVFTGTL 527
OY 420 RYFRALSOETGETLWOTRLATVASSQALSYVDGQVYAAIAG--GVSYGSL 470
DB 528 GLKARNDSIGELMKRLPSGLVGYPMTEYERHGVQVYIAHMSGVGMGVGL 579

RESULT 12
ID 091935 PRELIMINARY: PRT: 599 AA.
AC 091935;
DT 01-OCT-2000 (TREMBLER, 15, Created)
DT 01-OCT-2000 (TREMBLER, 15, last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
GN MXAF.
OS Methylovorus sp. (strain SSI / DSM 11726).
OC Bacteria; Proteobacteria; Delta subdivision; Methylophilus group.
OC Methylovorus.
OX NCBI_TaxID=81683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI;
RA Kim Y.M.;
RT "Cloning and nucleotide sequence of mxar gene of Methylovorus sp.
RT strain SSI DSM11726." to the EMBL/Genbank/DBJ databases.
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF184915; AAD56237.2;
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
SQ SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

Query Match 15.3%; Score 386; DB 2; Length 599;
Best Local Similarity 23.5%; Pred. No. 3e-20;
Matches 137; Conservative 89; Mismatches 209; Indels 148; Gaps 20;

OY 4 TSLMASGALLALAAPAFQVTPYTDLLANPPAGEWISYGOENQENYRHSPLTOITTEV 63
DB 6 TATGFAVAGLALSAMLPVAAAADSLALGANP--NNWMTQGTGYGQHSRLSQTITGN 63
OY 64 VGOLOL---VMARGKQPGKVQYTPPLIHDSVYTL----- 93
DB 64 VKNLESGIGHSTGLTLGH--EGAPLIVGDMYINPFPNNFALNLAEPEKIYWOHKPRQ 122
OY 94 -ANIVA-----NGVIA-----GSTCOY 110
DB 123 DASVKAACCDIYNRGLAYDDGHIFKTQLDGLHVASDAKTGKELMKMENDCPNAGSITQ 182
OY 111 SPF-----GC-----FVSGHDSATGELMRNFI-----PRA 137
DB 183 APFAVAGVLYVCGSAGELGVGRHAYANVKTIGEQARKYVATGDEELGLADDFNSAN 242
OY 138 GEE--GDETMGNDYEARMWTCG--WGQITDPTVNLVHYGSTAVGASPASETQRTGTPG 192
DB 243 GQMGICGTWBE--GWKIGGTWGMWYAYDPRKLNLIYSGSNAPAWNMTAR--PGDNK 297
OY 193 YGTNTRFAPVPTGELVWRHQLPRDNDQECTEFEMVTVNDVOPSTEMGLQSLINPNA 252

```

Db 298 W-TMTIARDVDYGAAGKMGYKTPHDEMDAGVQNMILTDQAVNGKTY----- 344

QY 253 TGBRRLGVPCKTMTMOQDAETGEFLMARDTN-YQNMIESIDENGIYVNEADILKEI 311

Db 345 -----FLTHVHD-RNGIMTTLNRQTSIVQAKAYDPVAVNFKKVDLKTGIPVADPEFSTRM 399

QY 312 DVE-YDVCPTFLGGRDWPSPALNPDGSIYFIPILNNCYDMAVDOEFTSMOYNTSNVTK 370

Db 400 DHKGTNLCPSAMGFHNGGLADYDPDRTYFGLNHCMDMEFPMIPYRAGOFEVGATLAM 459

QY 371 LP-----FGKMDIGRIDIDISTRTIMSVERAANYSPVLTSGGVLFGNGTDRYFRALS 426

Db 460 YPGNGPFTKEMQOVLAMGVGEEKYKMTKEKFSYWGGLATKGLVFTYTLNDOIRKALD 519

QY 427 QETGETLMQTRLATVASGAISYEVDGMQYVAIAGGVSYSG 469

Db 520 KTNKGTETKFKMPSGAIGAPMSYAYKGYIA-----TNGVG 557

RESULT 13

Q9AGW3 PRELIMINARY; PRT; 623 AA.

AC Q9AGW3: 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

OS 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.

OC Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=86174;

RN [1]

RP SEQUENCE FROM N.A.

RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.

RT "Characterization of the expression of two distinct alcohol

butanovora."

RL dehydrogenases involved in butane metabolism in Pseudomonas

submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF326086; AAK15506.1;

DR HSSP; Q944J7; 1FLG.

DR InterPro; IPR002372; Bac_POO_repeat.

DR Pfam; PF01011; Bacterial_POO; 4.

KW Signal.

FT SIGNAL.

SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 14.7%; Score 371; DB 2; Length 623;

Best Local Similarity 23.3%; Pred. No. 4e-19;

Matches 147; Conservative 75; Mismatches 203; Indels 206; Gaps 24;

QY 2 KPTSLMSAGALALLAFAFAQVPTVDLTLAN--PRAGEISISGQENYRHSPLTOI 59

Db 7 KPALRAIVVAATAALSLPA-AAVTDYTMEDIANDKTKTGDLVLTGLKAQRHSPKAI 65

QY 60 TTEVNGQLOLWVA-----RGMQPKQVYPLIHDGVMY----- 92

Db 66 NTDVNAVLVPMASFSGEKGQOE---AQV--LVHDDVIYATASYSIFAIDARSKKL 120

QY 93 -----LANIVANGYIVASTCOYSPFCFSGSDATGELKRNIF----- 133

Db 121 WEINARLPDDIRPCDVNRGAIAITGVFTGLDAAVNLDRKTKGVKVRKFGDHKYG 180

QY 134 -----IPRAGEEDDET----- 145

Db 181 YTWGAPYIKOKSGRILLVHGSSGDFGVVGLFARDPDTGEVWAPRWEGHGRIN 240

QY 146 GND-----YEARWMTGA---WQIITDPTNLVHGSTAVGPA 180

Db 241 GKDSFTGPKAPSWPDDPNSPTGKVEA--WSGGGAPWOTASFDVNNNVYIGAGNPAPW 299

QY 181 SETORGTG-----GLYGTNFAVRPDTGEIVWRHQTLPDWMDOECFEMVNTN-- 232

Db 300 NTKRTAPGDDPRRWDSLF-TSGQAVVASTGELKGFYQIIPNDANDPSGNSNVLETK 358

QY 233 -----VDVOPSTEMEGLOSIN-----PNAATGERRLVTCVCKTGTMMO--F 272

Db 359 DPKTKKNVNSAHADHNGFFFTVDRMLAKAGGPNKFTS---LIGAMPVDDITVASGF 415

QY 273 DAETGEFLMARDTNQNMIESIDENGIYVNEADILKELDVEXDVCPFLGGRDWPSPAL 332

Db 416 DLTKGRPI-EKDRPPQPEKAGDKESIFV-----PPFLGGNNHPSY 459

QY 333 NPDSGIYFIPILNNCYDMAVDOEFTSMOYNTSNV-----KLPPG-----KDMIG 379

Db 460 SPDTGLFYLPAHHMAMD-----YHMENTYKKSAYLQGGFRINLEDDHVG 506

QY 380 RIDAIDISTRTIMSVERAANYSPVLTSGGVLFGNGTDRYFRALSOETGETLMQTRLA 439

Db 507 IIRAIIDPSPARSLAGQGVRA-VAGILTITAGWVFTGSDGYLAKAPDAKNGKELMKRQTS 565

QY 440 TVASGAISYEVDGMQYVAIAGGVSYSG 470

Db 566 SGVSVVPTWEMDEQYVAIAGS-----YGVAV 593

RESULT 14

Q934G0 PRELIMINARY; PRT; 695 AA.

AC Q934G0: 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

GN LUPANINE HYDROXYLASE PRECURSOR.

OS Pseudomonas sp. DH2001.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=163360;

RN [1]

RP SEQUENCE FROM N.A.

RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,

RA Rogozinski J.

RT "Cloning, sequencing and analysis of the gene for lupanine

hydroxylase, a quinoxalochrome c from a Pseudomonas sp."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ318095; CAC67410.1;

KW Signal.

FT SIGNAL.

FT CHAIN

SEQUENCE 695 AA; 74956 MW; 9096C6387E457FE0 CRC64;

Query Match 14.4%; Score 364.5; DB 2; Length 695;

Best Local Similarity 24.2%; Pred. No. 1.4e-18;

Matches 130; Conservative 67; Mismatches 198; Indels 143; Gaps 19;

QY 38 AGEISISGQENYRHSPLTOITTEVNGQLOLVVARMOGKQYV--TPLIHGYMYL--- 93

Db 36 SGWMSLLGGGNDQHYFSALKDNKNSNVKNLGLSWFTMEKGDGLVNPVLADGVYIGGP 95

QY 94 -ANIVANGV-----IVAGSTCO----- 109

Db 96 PKGIYANDLKTGKNLMTVTPVOYDKDTSMTGFWGTHVNRGLAVDDNVYISGYLLAV 155

QY 110 -----YSPFG-----FVS-----GH-----DSATGEL 128

Db 156 SRTTKLITWSSQSDPKKMOAITGAPRVGCKVFIGNASGDRGGRHLDAFDKTKGKHL 215

QY 129 WRNFIIPRAGE-----EGDEFWGDYDEARHM-----TGAWGQITVDPTVNLVHY 172

Db 216 WREYTMP--GDPSKPFENDLLAKASKTWGTDY---WKYTKGVSPWAIITVDEASDTLYF 270

QY 173 GSTAVGASSETQGTGGLYGTNTRFAVRPDTGEIVWRHQTLPDWMDOECFEMVNTN 232

Db 271 GTDPSFWSPAQARAPDAGDELFSHIIAVDASTGAYKWHQFTQNDGSNNSATWHTMLAD 330

QY 233 VDVOPSTEMEGLSINPNATGRRRLVLTGVPCKTGMQFADTGEFLMARTNYONMIE 292
 Db 331 LPV-----EGV-----SKRVVMTAP-KNGEYVLDASTKFTFSADHYEVPNTK 373
 QY 293 SID-ENGIVYNDALIKELDYEDV-CPTFLGGRDWPSSAALNPSGIVYFLNNVCYDM 350
 Db 374 GLDPKTRPFPSEANWVEPGEKMTIPLPGDVGHHWEMAMYNPELRTVYIPSTLVPVTV 433
 QY 351 MA-VDOFTSMQY-----NTSNVTKLPKPKDMIGRIDAIDISTGRTLSVERAANSP 404
 Db 434 VASMDTGEILDYXXGMRPDATIKTO-----GDLVAMDPDLKKEKRAKRS.LPVNG 485
 QY 405 VLSGGGVLNNGSTDRFRALSOEGETIMQTRLATVASGAQISYEVDMQVVALAG 462
 Db 486 VLATAGGLVFOGTGDGHEAFDANTGKLSFHVGSIIAAPTVEVDGDTLYIVASG 543

RESULT 15
 Q59540
 ID Q59540 PRELIMINARY; PRT; 573 AA.
 AC Q59540;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
 DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)
 DE MEFHANOI. DEHYDROGENASE HEAVY SUBUNIT (EC 1.1.99.8).
 OS Methylophilus methylotrophus (Bacterium M3A1).
 OC Bacteria; Proteobacteria; Beta subdivision; Methylophilus group;
 OC Methylophilus.
 NCBI_Taxid=17;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M3A1:
 RA Xia Z.-X., Bai W.-W., Zhang Y.-F., He Y.-N., White S.A., Boyd G.D.,
 Matthews F.S.;
 RT "Determination of the Gene Sequence and the Three-dimensional
 structure at 2.4 A resolution of Methanol Dehydrogenase from
 Methylophilus methylotrophus.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U01040; AAA83765.1; .
 DR HSSP; P38539; AAAH.
 DR InterPro; IPR001479; Bac_PQQ.
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ; 7.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 573 AA; 62635 MW; A06C9B3091BB8F0C CRC64;

Query Match 14.2%; Score 359.5; DB 2; Length 573;
 Best Local Similarity 23.0%; Pred. No. 2.5e-18;
 Matches 127; Conservative 75; Mismatches 197; Indels 153; Gaps 20;

QY 38 AGEMISTGQNGENRHSPLTITENYGOLOLV-ARGM----- 75
 Db 13 AGAMPATGAGYYSQNSPLAQINKSNKVAANSFSTGVLNHGEAPLYIGDMATVHSA 72
 QY 76 -----QPK-----VOVPLIHGVMYLANIV-----ANGVIVA- 104
 Db 73 FPNNTYALNMDPKIYVQHKPKODASTKAMVCDVDRGLAYGAGIYKKQANGHLAL 132
 QY 105 -----GSTQYSPF-----GC-----FVSGHDSATGGEIM 129
 Db 133 DAKTKINMEVEVCDPKYGTSTLQAPFVAKDVLMMGSGAELGVRGAVNAEDLKTGLKW 192
 QY 130 R-----NYFIPRAGE--EGDETWNDEYEAAMTGA--WGQITVDPVTN 168
 Db 193 RAFPATGSDSVRLAKDFNSANPHIGQFGLGKTIWEGD--AKKIGGGINMGWYAYDPKLN 249
 QY 169 LVHYGSTAVGPASETQSTPGG-----TLVGTNTRFAVRPDTGEIIVRHQTLPRDNDOE 223
 Db 250 LFYYSGGNAPAPNMEIMR--PGDNKWTNTIWRDL-----DTGMAKMGYQKTPHDEWFA 301

QY 224 CTEEMAVTVNDVOPSTEMEGLSINPNATGRRRLVLTGVPCKTGMQFADTGEFLMAR 283
 Db 302 GVNQAVLTLD-----QPVN-----GKMTPLSHIDRNGILTYLNRNGNLVYAE 344
 QY 284 DTN-YQNMIESIDENGIVYNDALIKELDYEDV-CPTFLGGRDWPSSAALNPSGIVYFL 341
 Db 345 KVDPAVWVEFKVDLKTGTPVDRDEFATRMDHKGTNICSAMGFHMQVDSIDPESRTIYA 404
 QY 342 PLNNVCYDMAVDOFTSMQY-----NTSNVTKLP-----PKDMIGRIDAIDISTGRTLSVER 397
 Db 405 GLNHICDMWEPFMLPFRAGQFFVGAATLAMYPGNGPTKKEMQIRAFDLTGKAKRTWE 464
 QY 398 AAANYSPVLSTGGGVLPNGSTDRFRALSOEGETIMQTRLATVASGAQISYEVDMQV 457
 Db 465 KFAAMGSTLYTKGLVYATLDDYLKALDNKDKELMNFKMPGSGIGSPMTYSFKQOYI 524
 QY 458 ALAGGYSYSGG 469
 Db 525 -----GSMYGVG 531

Search completed: May 24, 2002, 10:30:15
 Job time: 648 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:16 ; Search time 101.42 seconds
(without alignments)
549,784 Million cell updates/sec

Title: WALICK-934-125.PEP
Perfect score: 2689
Sequence: 1 MKPSTLWASAGALALAAP.....GMOYVIAAGGVSGSGINS 502

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2652	98.6	579	19	AAW37873
2	2327	86.5	579	19	AAW37874
3	2304.5	85.7	579	19	AAW37876
4	2233.5	83.1	578	19	AAW37875
5	476	17.7	754	21	AAW35987
6	474.5	17.6	738	13	AAW20192
7	474	17.6	742	11	AAW52335
8	473.5	17.6	738	12	AAW33993
9	307.5	11.4	740	20	AAW5019
10	153	5.7	443	22	ABG24430
11	140.5	5.2	715	22	AAW81738

12	140	5.2	910	22	AAW33007	S. epidermidis ope
13	139.5	5.2	696	22	AAW32914	S. epidermidis ope
14	136.5	5.1	470	22	AAW32701	S. epidermidis ope
15	129	4.8	948	22	ABG25909	Novel human diagno
16	129	4.8	1510	22	ABG21573	Novel human diagno
17	126.5	4.7	593	11	AAW07014	Protein G variant
18	126.5	4.7	1016	22	AAW34349	Staphylococcus aur
19	125.5	4.7	824	18	AAW09614	Sequence of polype
20	125.5	4.7	269	11	AAW07006	Pathogenic Staphyl
21	119	4.4	269	15	AAW3295	Igg-binding Strept
22	119	4.4	269	15	AAW3295	Streptococcus prot
23	118.5	4.4	593	15	AAW62944	Protein G variant.
24	118	4.4	594	10	AAW10005	Streptococcus GX78
25	116	4.3	594	12	AAW10005	S. epidermidis ope
26	114.5	4.2	404	22	AAW82772	S. epidermidis ope
27	112.5	4.2	514	22	AAW95225	HPV6 mutant L1 pro
28	110.5	4.1	278	22	AAW82625	HPV6 mutant L1 pro
29	110.5	4.1	501	18	AAW10645	Expression vector
30	109.5	4.1	228	21	AAW10643	HPV6 mutant L1 pr
31	109.5	4.1	501	18	AAW10644	HPV11 mutant L1 pr
32	108.5	4.0	501	18	AAW26108	HPV11 mutant L1 pr
33	108.5	4.0	501	18	AAW26109	Human PRO4340 poly
34	108.5	4.0	530	22	AAW12235	Human PRO4340 poly
35	107.5	4.0	770	22	AAW30824	HPV11 mutant L1 pr
36	107.5	4.0	998	10	AAW33076	Amino acid sequenc
37	107.5	4.0	998	10	AAW33621	Sequence of the an
38	107.5	4.0	998	18	AAW33621	Elemeria tenella 45
39	106	3.9	649	22	ABW6538	Drosophila melanog
40	106	3.9	2659	22	ABW68426	Drosophila melanog
41	105.5	3.9	429	22	AAW66542	Human interferon-a
42	105.5	3.9	429	22	AAW66542	Human protein sequ
43	105.5	3.9	469	22	AAW6578	Human prostate can
44	105.5	3.9	501	21	AAW99981	HPV 6 L1 protein m
45	105.5	3.9	996	18	AAW33624	Elemeria tenella 45

ALIGNMENTS

RESULT 1	AAW37873	AAW37873 standard; Protein: 579 AA.
XX	AAW37873	
AC	AAW37873	
DT	10-AUG-1998	(first entry)
XX		
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX		
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulonic acid; L-ascorbic acid; Inhibition.	
OS	Glucanobacter oxydans.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..23
FT	Protein	/note= "signal peptide"
FT		24..579
FT		/note= "mature protein"
XX	EP832974-A2.	
PN		
XX	01-APR-1998.	
PD		
XX	11-SEP-1997.	97EP-0115801.
PF		
XX	19-SEP-1996.	96EP-0115001.
PR		
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
PA		
XX	A. alioceitigenes me	
XX	Sorbitol dehydroge	
XX	Novel human diagno	
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
XX		

DR WPI: 1998-195228/18.
DR N-PSDB: AAV29051.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde
dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1: Pages 35-37; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans
alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
or recombinant organisms can be used to convert suitable substrates
to aldehydes, ketones or carboxylic acids, especially to convert
L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
converted to L-ascorbic acid by standard procedures. The derivatives
of ADH enzymes have desired substrate specificity, higher affinity
to a substrate, lower affinity to an inhibitory compound, higher
stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA:

Query Match 98.6%; Score 2652; DB 19; Length 579;
Best Local Similarity 90.3%; Pred. No. 7.8e-227;
Matches 502; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 1 MKPTSLIMASAGALALAPARAOVTPVDELANPPAGEMISYQONENYRHSPLTQIT 60
DB 1 mkptslimwasagalalapaatavtptdelanppagewisgqenyrhsplqt 60
QY 61 TENVGOLQIWMARGMPKQKVOVTPPLIHGVMYLANPGDVIAIDAKTGDIWHRQLPN 120
DB 61 tenvgqlqlwargmpkqvtpplihgvmylanpgdvialdaiktgdilwehrrqlpn 120
QY 121 IATLN-----I 126
DB 121 iatlnsfgeptlrmalygttnvfyvwdnhlvaldtatgvtfdvdrggedmwmsnsgpi 180
QY 127 VANGYIVAGSTCOYSPFCEVSGHDSATGELMKNFYIPRAGEGDETMGNDYKRWMTG 186
DB 181 vanygiavagstcgyfpfcvsgshdsatgelmrnyfipragegdetmgnndykarwmtg 240
QY 187 AMGQITDPVTNLVHGSTAVGPASEFQRTGCTIGNTNFAVRPDTGELWHRQLP 246
DB 241 awgqitdpytnlvhgstavgpasefqrctgctigntnfavrpdtgelwvhrqtlp 300
QY 247 RDNDQCTPEEMAVTNDVOPSTEMEGIQSINPNATGERRVLTGPCKTGEMMQPDAET 306
DB 301 rdndqctpeemvtnvdvqstemeqigslnpnaatgerrrvltgvpcktgemmqidaet 360
QY 307 GEPLMARDTNYQNMIESIDENGIVTVNBDALIKELDEYEDVCPFLGGRMPSAALNPS 366
DB 361 geflwardtnyqmiesidengivtvebdalikeldeydvcpflggrdmpsaalnpds 420
QY 367 GIEFELNNVCYDMAVDOETSMQVNTSNVTKLPKQKIDIGRIDIDISTGRITLSWVE 426
DB 421 giefelnnvcydmavdoetsmqvntsnvtklpkqkidigriddistgritlswve 480
QY 427 RAAANTSPVLSGSGVLNFGCTDRFRALSGETETLMQRLATVASGAISYEVDMQY 486
DB 481 raaantspvlsrgsgvlnfgctdrfralsgetetlmqrlatvasgaaisyevdmqy 540
QY 487 VALIAGGSYSGGLNS 502
DB 541 valaggsvsgsglins 556

RESULT 2

AAW37874

XX AAW37874 standard; Protein; 579 AA.

AC AAW37874;

XX

DT 10-AUG-1998 (first entry)
XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
DE Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
XX aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
XX

Gluconobacter oxydans.

Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note="signal peptide"

FT Protein /note="mature protein"

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI: 1998-195228/18.

N-PSDB: AAV29052.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde

dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1: Pages 38-40; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans

alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes

or recombinant organisms can be used to convert suitable substrates

to aldehydes, ketones or carboxylic acids, especially to convert

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be

converted to L-ascorbic acid by standard procedures. The derivatives

of ADH enzymes have desired substrate specificity, higher affinity

to a substrate, lower affinity to an inhibitory compound, higher

stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA:

Query Match 86.5%; Score 2327; DB 19; Length 579;
Best Local Similarity 78.3%; Pred. No. 6.1e-198;
Matches 432; Conservative 32; Mismatches 34; Indels 54; Gaps 1;

QY 1 MKPTSLIMASAGALALAPARAOVTPVDELANPPAGEMISYQONENYRHSPLTQIT 60
DB 1 mkptslimwasagalalapaatavtptdelanppagewisgqenyrhsplqt 60
QY 61 TENVGOLQIWMARGMPKQKVOVTPPLIHGVMYLANPGDVIAIDAKTGDIWHRQLPN 120
DB 61 tenvgqlqlwargmpkqvtpplihgvmylanpgdvialdaiktgdilwehrrqlpn 120
QY 121 IATLN-----I 126
DB 121 iatlnsfgeptlrmalygttnvfyvwdnhlvaldtatgvtfdvdrggedmwmsnsgpi 180
QY 127 VANGYIVAGSTCOYSPFCEVSGHDSATGELMKNFYIPRAGEGDETMGNDYKRWMTG 186
DB 181 vanygiavagstcgyfpfcvsgshdsatgelmrnyfipragegdetmgnndykarwmtg 240
QY 187 AMGQITDPVTNLVHGSTAVGPASEFQRTGCTIGNTNFAVRPDTGELWHRQLP 246
DB 241 awgqitdpytnlvhgstavgpasefqrctgctigntnfavrpdtgelwvhrqtlp 300

XX 01-APR-1998.
 XX 11-SEP-1997; 97EP-0115801.
 XX 19-SEP-1996; 96EP-0115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N.
 XX WPI; 1998-195228/18.
 XX N-PSDB; AAV29053.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1; Pages 41-43; 59pp; English.
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX Sequence 578 AA;

Query Match 83.18; Score 2233.5; DB 19; Length 578;
 Best Local Similarity 75.18; Pred. No. 1.2e-189;
 Matches 417; Conservative 37; Mismatches 46; Indels 55; Gaps 2;

QY 1 MKPTSLMASAGALALALAFPAFAOVPTDELLANPAGENISTGONENRHSPLQITM 60
 DB 1 mktllqssaallvltipalag-taltdemlanpaggewlnyqngenyrsplqt 59
 QY 61 TENVGQQLVWARGMOPGVQVTPILHDCVWLANPGDVQADIAKGDILMHRQQLN 120
 DB 60 adnygqqlvwarameagklytvlphdgvmylanpgdvqadialdewhrrqlpn 119
 QY 121 IATLN-----I 126
 DB 120 iatlnsfgeptgmalylgtnyfyvwdhvaldtsqgvvfvdrggqtdmvsnsqpl 179
 QY 127 VANGVIVAGSTQYSPFCFVSGHDSATGEELMRNFITPRAGEGDETMGNDEARMMTG 186
 DB 180 vanyvfvagstcqsypfcitvsgndsatgeelwrmnfipragegdetwngndearmtg 239
 QY 187 AAGQITDYPVNLVHYSTAVGPASETORCTPGCTLGTNTFEAVRPTGELVRRHOTLP 246
 DB 240 vwgqitdypvnglvhytsavgpaaefqgtvsgmynctfavrptgclvrrhotlp 299
 QY 247 RDNMDOECTEMWTVNDVOPSTEMEGLSINPMNATSGRRVLTGVPCKTGMQFAET 306
 DB 300 rdndogectemwvndvopstemeqlhainpdaatgerrvltgvpckngtmqfaet 359
 QY 307 GEFLMARDIYONMISIDENGLIVYNEDAILKELDYEDVCPFTLGRMPASALNPDS 366
 DB 360 geflwardisynmliesvdpdglvynedlvvtelevayelcplfigrdmpsaalnpdt 419
 QY 367 GTFITPLNNVCYDMAVDOEFTSMOVNTSVTKLPKGMIGRIDALISTGRTLMSEV 426
 DB 420 gtyfipilnnacsgmlavdeftsmovntsvtklpkigmigridalistsgrtlmsev 479
 QY 427 RAANANSPVLSGGVLENGSTDRFRALSOETGETLMQRLATVASGAISYEDMGQY 486
 DB 480 ryasnospvlsrggvlengstdrfralsetgetlmqrlatvasgaaisyedmgvy 539

QY 487 VALAGGVSYSGGLN 501
 DB 540 valagrvlcsygsnln 554

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.

AC AAB35987;

PT 01-MAR-2001 (first entry)

DE Sorbitol dehydrogenase subunit 1 amino acid sequence.

XX Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 XX L-sorbose production; 2-keto-L-gulonic acid.

OS Gluconobacter oxydans.

PN MO200065066-A1.

PD 02-NOV-2000.

XX 23-APR-1999; 99WO-IB00736.

PR 23-APR-1999; 99WO-IB00736.

PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.

PI Choi E, Rhee S, Lee E;

DR WPI; 2000-687351/67.
 DR N-PSDB; AAC83153.

PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol

PS Claim 1; Fig 8; 96pp; English.

CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose and AAC83157 which contain the subunit
 CC production of 2-keto-L-gulonic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.

SO Sequence 754 AA;

Query Match 17.7%; Score 476; DB 21; Length 754;
 Best Local Similarity 26.3%; Pred. No. 3.1e-33;
 Matches 158; Conservative 80; Mismatches 230; Indels 132; Gaps 18;

QY 14 LALLAAPFAOVTPV-----IDELLANPAGENISTGONENRHSPLQITTEVNGQ 66

DB 17 lgcfaalaefcstspvalaedgtaltnadqhpqdmmsygrtysqiryspldqtikdaasn 76

QY 67 LQVWAKMOPGVQV-TPLIHGVWYLANPGDVQADIAKGDILMHRQQLN-NIA-- 122

DB 77 lqlwvhydltdtnrggeftplivdgymyatlmwskmkaalaaqtlwsvqpkvynladr 136

QY 123 ---TLNIVA---NGVIVAGS-----TCQTS 141

```

137 gcoctvngaaaywngkvfygftdgrllaldaktkglwsvytwpkkaqjghgrsyvtdga 196
142 P-----FGC--FVSGHDSATGEELMRNVPFRAGEGD----- 172
197 prlakgkvllnggaefgargftaydaetckmdwrfittvnpnhkpkdgaasddvmska 256
173 -ETWGDYDEARMWTG---AMGQITPDVYTNLVHKGSTAVGPASTQGTGCTLXGNR 228
257 ytpwagagawkgqgsggtvwsdllydpytdlylvgngspwnyktfseggmnlfigsl 316
229 FAVRPDTEGIWRHQTLPDRNDQCEPEMWTNVDOVSTEMEGLOSINFNATGE-RR 287
317 valnptdtkyvvwhfgetpmdqwtysvqgimalampv-----ngemrh 359
OY 288 VLVGVPCKTGMQFPAETGEELMARDTYQNMIESIDE-NGIYVNDALIKELDEYD 346
Db 360 vlvnap-kgnfilyldaktkgkfskpylyenwangldpvtgrnynpdalwtlmgkpw 418
OY 347 VCPFLGGRDWPASALNPDSGITFPLNNVCY-----DMAVDQETSMDEV 393
Db 419 gipdligmhaamayspqtkllyipagvpyvdpqkgfkhahdswmlgldmkgll 478
OY 394 NTSN---VTKLPCKMIGRIDAIDISTGRILMSVERAANSPVLSTGGGVLENGGD 449
Db 479 ddndpqhkdakqfklkdlkgwlvawdpqkgaafvtdhkgpwnngllacagvllfglan 538
OY 450 RYFRALSOETGLMOTRLATVASGOAISYEVDGMQVAT---AG-----GVSYSG 499
Db 539 gefhaydattgkdlftrpagsallappvtylangkqyvaevwgsllypflgvyarts 598

RESULT 6
AAR20192 standard; protein: 738 AA.
ID AAR20192
XX AAR20192;
AC 16-APR-1992 (first entry)
DT 16-APR-1992
XX ADH complex protein (mol.wt. 72.000).
DE ADH complex protein (mol.wt. 72.000).
XX Alcohol dehydrogenase; acetic acid; fermentation.
XX Acetobacter alioacetilgenes NH-24.
XX JP03266988-A.
XX 27-NOV-1991.
XX 26-MAR-1990; 90JP-0073440.
XX 26-FEB-1990; 90JP-0042301.
XX 26-MAR-1990; 90JP-0073440.
XX (NAKA-) NAKANO SUTEN KK.
XX WPI: 1992-019325/03.
XX N-PSDB: AAQ20383.
XX Alcohol dehydrogenase complex structural gene - used in plasmid
XX PT and enhancing efficiency of acetic acid fermentation for
XX PT transformed acetic acid bacteria
XX PS Disclosure; fig 3(1-3); 21pp; Japanese.
XX Acetobacter transformed with the sequence encoding this protein can
XX enhance the efficiency of acetic acid fermentation. The ADH complex
XX can be easily extracted from the bacteria and purified and it can be
XX used for the determination of an alcohol.
XX See also AAQ20383-84, and -86-88.
XX SQ Sequence 738 AA;

```

```

Query Match 17.6%; Score 474.5; DB 13; Length 738;
Best Local Similarity 25.6%; Pred. No. 4e-33;
Matches 151; Conservative 77; Mismatches 229; Indels 133; Gaps 15;

OY 10 SAGLALLAAPAFQVTPYVDE-----LIAPPAGEMISYGNQENRHSPLTGIT 61
Db 17 tagtcaallsyvatmasaddggatgatcaalhhadhpymwlytgsdqtyspldqtnr 76
OY 62 ENVGQILVWARKQPEKVOY-TPLIHDGVWYLANPGVDIAIDAKTGLIWEHRQLP 119
Db 77 snvgnlkawyldldtnrggeqplvdygwyattwsmmkvadaatgkllwsyqprvpy 136
OY 120 NTA-----TN----- 125
Db 137 nhadggcoctvngaaaywngkvfygftdgrllaldaktkglwsvntlppeaejgkqrsy 196
OY 126 -----IVANGVIYAGSTCOYSPRC--FVSGHDSATGEELMRNVPFRAGEGD----- 172
Db 197 tvdgaprlakgrvllgn--ggsefgargvsaftaetgkvdwrfittvnpknepdaads 254
OY 173 -----ETWGDYDEARMWTGA-----WGQITPDVYTNLVHKGSTAVGPASTQ 215
Db 255 vlmnkayqlws-----ptgawtrgsggtvwsdllydpytdlylvgngspwnykyr 307
OY 216 GTPGGTLVGTNTRFVRPDPDTEGIWRHQTLPDRNDQCEPEMWTNVDOVSTEMEGIQ 275
Db 308 segkgdnlfigslvalkpetgeyvvwhfgetpmdqwtfdtsdqglltdlpl----- 357
OY 276 SINPNATGERRVLGVPCKTGMQFPAETGEELMARDTYQNMISID-ENGIVTNE 334
Db 358 -----ngeltrivharhknngfilyldaktkgkfskpylyenwangldpvtgrnyn 410
OY 335 DAIKELDEYDVCPTFLGGRDWPASALNPDSGITFPLNNVCYDMAVDQET-SMDVY 393
Db 411 dalyltqkewyigpdlqgmhaamayspqtkllyipagvpyvdpqkgfkhahdswmlgldmkgll 470
OY 394 NTS---NWTKLPPG-----KDMTGRIDAIDISTGRILMSVERAANSPVLSTGGGV 442
Db 471 nlglmdmkvgipdspeakqafvkdikgwlvawdpqkgaawtvtdhkgpwnngllacagvllfgldl 530
OY 443 LFNQGTDRYFRALSOETGLMOTRLATVASGOAISYEVDGMQVATAG 492
Db 531 lfglangefhaydatngsdlnfaadsgliappvtylangkqyvaevwg 598

RESULT 7
AAR05235 standard; protein: 742 AA.
ID AAR05235
XX AAR05235;
AC 04-AUG-1990 (first entry)
DT 04-AUG-1990
XX Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
XX DE 1-2229 of alcohol dehydrogenase (ADH) gene.
XX Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
XX Acetobacter aceti K1006 (FERM-7528); plasmid pAA721.
XX Acetobacter aceti strain K1006 (FERM-7528).
XX JP02000452-A.
XX 05-JAN-1990.
XX 30-OCT-1987; 87JP-0273190.
XX 30-OCT-1987; 87JP-0273190, JP-075069.
XX (QPPP ) QP CORP (QPPJ-).
XX PA
XX

```

DR WPI; 1990-047990/07.
 DR N-PSDB; AAQ91811.
 XX
 PT Cloning DNA, plasmid and microbe contg. it -
 PT contains alcohol dehydrogenase gene derived from Acetobacter
 PT aceti K10006 strain
 XX
 PS Disclosure: Fig 4; 8pp; Japanese.
 XX
 CC Also new are a recombinant plasmid contg. its encoding DNA, and a
 CC microorganism transformed with the plasmid.
 CC A DNA fragment was obtd. from A. aceti
 CC strain K1006 and was transferred to an ADH-defective strain. The
 CC resulting plasmid, PA4721, contg. its gene was inserted into A.
 CC pasteurianus strain IFO 3191 by the triparent method. Its gene is
 CC useful for improving Acetobacter culturing.
 XX
 SQ Sequence 742 AA;

Query Match 17.6%; Score 474; DB 11; Length 742;
 Best Local Similarity 26.0%; Pred. No. 4.5e-33;
 Matches 160; Conservative 87; Mismatches 202; Indels 166; Gaps 23;

OY 11 AGALALIAFAFAQVPTVDEL--ANPPAGEWISYCONENRHSPLTOITTEWGOLO 68
 DB 23 aaalpaayapaaraagagntgeaithaddipenwisygrtyseqryspldqnsrvnvgdk 82
 OY 69 LVNAROMOGKVO-VPLIHGVMYLANPGDVIAIDAKTGDLIWEHRROLN- 122
 DB 83 llyytlidnrgaeatpivdgimayattwskmealaatgkllwydqpknaladkgo 142
 OY 123 --TLN----- 126
 DB 143 cdtvngagayngkvfgtfdgrlvaadatkqkhwavntlpadaalqkqrsyvtgavvr 202
 OY 127 VANGVIVAGSTQYISPEFC--FVSGHDSANGELMRYFIPRAGEED- 172
 DB 203 vakjytlvlgng--ggaefagrfvaafdaegkikwrtfvpknknepdhaadnllmka 260
 OY 173 -ETWGNDEYARMWT-----GAMGOITDPTNIVHGSTAVGASASTONGTREGTLY- 223
 DB 261 yktwpr--kgaavvrqggsgtwdslvydpsdily--lavg-----ngspwmykyrse 309
 OY 224 --GTV-----TRFAVRPTGELVWRHQTLPEDNWDCEPTENWVTWVQPTSTEMGLQSI 277
 DB 310 gisgnlflgslvalkpetgeyvwahfagatpmdgwdytsvgqintldmpvk----- 358
 OY 278 NPNAATGE-RRVLTVGVCCKTGMOPDAENGELIAROTNYOMTESID-----EN 327
 DB 359 -----gemhvlvhar-kngffvldaktgeflsgknyvqynwngldp1tgrpmynd 411
 OY 328 GIYVUNEDALIKELDYDYVDCPEFLGGRDPSALNPDGFIYFLNNVVC----- 378
 DB 412 gilytling-----kfwygl--pqp1gahfmmamayspklthvypahqiprpykngvggf 463
 OY 379 -----DMMAVDOETSMDEVNTSNVTKLPKPKDMIGRIDAIISIGRTLWSEARAANYSP 434
 DB 464 kphadeswvngldmtknglpdipe-artayikdlngwllawbrvkmecvkwidhngpvg 522
 OY 435 VISTGGGVFNGCTDRYRFAISQETGETLMOJRLATVAAAGAISEYEDGAGQYVA----- 488
 DB 523 llatcgddllftglangefhaydatnsgdlykfdagsgliapmptysvngkqyvaavewg 582
 OY 489 ----IAGGVSYSYSGC 499
 DB 583 gilypsmgvgvgtsg 597

RESULT 8
 AAR13993
 ID AAR13993 standard; Protein; 738 AA.
 XX

AC AAR13993;
 XX
 DT 09-DEC-1991 (first entry)
 XX
 DE A.alcoetigenes membrane-bound ADH 72kd sub-unit.
 XX
 KM alcohol dehydrogenase complex; carboxylic acid production.
 XX
 OS Acetobacter alcoetigenes.
 XX
 PN EP448969-A.
 XX
 PD 02-OCT-1991.
 XX
 PF 26-FEB-1991; 91EP-0102793.
 XX
 PR 26-MAR-1990; 90JP-0073440.
 XX
 PR 26-FEB-1990; 90JP-0042391.
 XX
 PA (NAKA-) NAKANO VINEGAR KK.
 XX
 PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;
 PI Kawamura Y;
 XX
 DR WPI; 1991-289462/40.
 XX
 DR N-PSDB; AAQ13580.
 XX
 PT Gene for membrane-bound alcohol dehydrogenase complex - obtd.
 PT from Acetobacter alcoetigenes, used for prodn. of enzyme for
 PT converting alcohol to acid
 XX
 PS Disclosure; Fig 3; 36pp; English.

CC Total DNA was prepared from A.alcoetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
 CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. alcoetigenes (see AAQ13582-013584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 XX
 SQ Sequence 738 AA;

Query Match 17.6%; Score 473.5; DB 12; Length 738;
 Best Local Similarity 25.6%; Pred. No. 4.9e-33;
 Matches 151; Conservative 76; Mismatches 230; Indels 133; Gaps 15;

OY 10 SAGLALIALAFAFAQVPTVDEL-----LANPPAGEWISYCONENRHSPLTOITT 61
 DB 17 tagtcaallssyatlmasaddggatgatgeaithaddipgnmmtlygrtyseqryspldqinr 76
 OY 62 ENWGQLOLVNAROMOGKVOV-TPLIHGVMYLANPGDVIAIDAKTGDLIWEHRROLN- 119
 DB 77 snvgnlklaylidlndnrgsggtprlvldgvmayattwskmkaadaacgkllwsydpvpg 136
 OY 120 NIA-----TLN----- 125
 DB 137 nldakgcddtvtvngaaayngkvfgtfdgrlvaldaktgkhwavntlpdaaalqkqrsy 196
 OY 126 ----IVANGVIVAGSTQYISPEFC--FVSGHDSANGELMRYFIPRAGEED----- 172
 DB 197 tvdgapriakgrvlgng--ggsefgagrfvaafdaegkikwrtfvpknknepdhaads 254
 OY 173 -----ETWGNDEYARMWTGCA-----MGQITDPTNIVHGSTAVGASASTONGTREGTLY 215
 DB 255 vlmnkaygtws-----ptgawtrtggsggttwdslvydpsdilykfdagsgliapmptysvngkqyvaavewg 307

QY 216 GTPGTLGTNTFRANVPDGEIWMRHQTLPRDNMDCECTFEMAVTNDVOPSTEMEGLQ 275
 Db 308 segkxgdlflfysvalpfcgcyvwhfgetpmdqwtcdtsdgmldlpi----- 357
 QY 276 SINPMATGERRVLTVGPCKGTMMQFDATGEFLMARPTNQNMIESID-ENGIVTNE 334
 Db 338 -----ngertihvharhkgffylidantgeifsgkyvvnwasgldpkygpiy 410
 QY 335 DAILELDEVEDVCPFLTGSDMPSALNPDGIVFIPLNVCYDMMAVDOEFT-SMDVY 393
 Db 411 dalyltlgkewyqipgdlygmniafapktqlyvipaqvqpflytnvggftphpsdw 470
 QY 394 NNS---NWKLEPPG-----KDMGRIDAIDISTRTIWSVERAANSPVLSITGGV 442
 Db 471 nlgldmknkylpdspeakqatfvdkgdkyivawdpkqaeswrvdnhkypwnngllatgqd 530
 QY 443 LKNGTDRFRALSGEETLMQTRLATVASGQAISEYVDGMQVAVIAGG 492
 Db 531 lfiglangeifhaydatngsdlfhfaadsgllaprvtylangkyvavevg 580

RESULT 9

AAM95019 standard; Protein; 740 AA.
 ID AAM95019;

AA95019;

21-MAY-1999 (first entry)

Sorbitol dehydrogenase (SLDH) protein sequence.

Sorbitol dehydrogenase (SLDH); open reading frame; ORF2; L-sorbose;

D-sorbitol; vitamin C; enzyme.

Gluconobacter suboxydans.

Key Location/Qualifiers

Peptide 1..24

Protein /note="signal sequence"

25..740

/note="mature protein"

EP897984-A2.

24-FEB-1999.

13-AUG-1998; 98EP-0115231.

21-AUG-1997; 97EP-0114432.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

WPI: 1999-134646/12.

N-PSDB; AAX11501.

New D-sorbitol dehydrogenase gene and recombinant protein useful for production of L-sorbose, an intermediate in vitamin C production

Claim 1; Fig 3A-D; 39pp; English.

This represents a sorbitol dehydrogenase (SLDH) protein. The DNA encoding the SLDH enzyme also encodes an open reading frame (ORF2) product upstream of the SLDH open reading frame, needed for SLDH activity in vivo. Host cells transformed by a vector comprising the sorbitol dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for producing L-sorbose from D-sorbitol. L-sorbose is an important intermediate in vitamin C production.

Sequence 740 AA:

Query Match 11.4%; Score 307.5; DB 20; Length 740;
 Best Local Similarity 21.6%; Pred. No. 2,8e-18;
 Matches 148; Conservative 74; Mismatches 195; Indels 269; Gaps 33;
 QY 23 AQVTPVTELLANPAGEMISYGOENYRHSPLTQTITENVGQLQ--VMWGM--OPG 78
 Db 86 sqvpmappqgsanpargdwagvrdhbgtrypslsetlpnaslkvlatyhtgspyrp 145
 QY 79 KV-----QVTPV-LHDGWTLANPGDVGQADAKTGDLWEHR-----RQLPNIAT--L 124
 Db 146 qvkvwaetpklvkgdyltcsamndilk-lpataqkqirrvndvkyhsipytaackgv 204
 QY 125 NIYANGVIYASTCOYSPGCFVSGH-----DSATGEELMKNYIIPRAGE----- 169
 Db 205 tyftssvpeggqch-----nrllegtlamrlavdaetgd-----lcpnfiggqvnlm 254
 QY 170 -----EGDETW-----GNDYEARMTGMAN--- 188
 Db 255 qglgesvpgfvsmtappvlnvgvvvnhveidgqrtrwapsvirydaesgkfvaawdn 314
 QY 189 -----GOITDPVETNLVHYGSTAVGPASETQKGTPEGTLXGTNFR-----FAVRPD 234
 Db 315 nsgrsqpay-rvtvltaverripqlpdrgrg-----gsrlrpdmsaadyalsds 366
 QY 235 -----TGEIWMRHQTLPRDNMDCECTFEMAVTNDVOPSTEMEGLQSIINP 279
 Db 367 aenkssavvaldvkltgspvrvftqahkxdwdydisgatl-----mdmpg-----p 413
 QY 280 NNAIGERRVLTVGPCKGTMMQFDATGEFLMARPTNQNMIESIDEN-----GIV--- 330
 Db 414 dqqlvpalim-pkrgqtfvlrrtqk-----pilhveirpapsppvlpqdp 459
 QY 331 -----TVNEDAT-----LKELD-----VEY 345
 Db 460 rslpqpwsvqmpalrtpdketdmwgmapiqlfclrkfranyvgeftppsvdkpwy 519
 QY 346 DVCPFLTGSDMPSALNPDGIVFIPLN-NCVCDMAVDOEFTSM----- 390
 Db 520 ---pyngsgdsgmsydpqsgllaanwltpmydqqlvtrkadsiglmipdnfknpg 576
 QY 391 -----DYNTSNVETKLPKGMIRIDAIDISFG-RTIMS--- 424
 Db 577 ggaengamdgiprylvtvprfwdqy-tgmncnrrp-----ygmilcdmkhkgkvlwqhl 631
 QY 425 -----VERAANYSPVLSTGGGVLFNG-GTDRYFRALSGEETLMQTR 467
 Db 632 gtatangpwlpqglpweigtprnngsvvtggllifgaatngiraidentgkvwsav 691
 QY 468 LATVASGQAISEYVDGMQVAVIAGG 493
 Db 692 lpggganpmtlyeanghqvaylmaag 717

RESULT 10

ABG24430 standard; Protein; 443 AA.

ABG24430;

18-FEB-2002 (first entry)

Novel human diagnostic protein #24421.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

Human; chromosome mapping; gene mapping; gene therapy; forensic;

Food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

XX	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
PR	
XX	

PI Drmanac RT, Liu C, Tang YT;
XX

PT New isolated polynucleotide a
PT diagnostics for

PT disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID No 54789; 103pp; English

the invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC disorders involving aberrant (II). (I) and (II) are useful for treating CC the polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AB600010-AB630377 represent novel human CC diagnostic amino acid sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at http://wipo.int/pub/published_pct_sequences. CC

Query Match	5.78;	Score 153;	DB 22;	Length 443;
Best Local Similarity	30.8%;	Pred. NO. 7e-05;		
Matches 44;	Conservative	33.14%		

QY	6	LIMS-----	AGALTLAAPAFAYPTVDELLANPPAGEWTSQONQENRSHPTQ	58	
Db	168	ltwagfdnpelngttsadatpaea	tspsvadq-----	dwpaygrngqgrfeplyq	218
QY	59	ITTENGGQLOLWYA-----	RGMPGRV--QVPLIHGVMTYANGDYIAIDAKTQD	109	
Db	219	lnadnvnhlkqavfrftgqdvkqndpgstlnevepikvdytltahqrlfaldaa	sgk-----		
QY	110	LIEHHRROL-PINATLNIYANGV	131		
		: : :			
Db	279	ekhygpekltnesfqhvtlrcgv	301		

RESULT	11
AA81738	
ID	AA81738 standard; Protein; 715 AA.
XX	

AC	AAG81738;
XX	
DT	03-SEP-2001 (first entry)
XX	

Accession	Open reading frame	protein sequence	SEQ ID NO:570
XX	Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination; endocarditis.		
KW			
KW			

XX Staphylococcus epidermidis.
OS
XX
PN WO200134809-A2.
XY

PD 17-MAY-2001
XX

09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.
XX

PA (GLAX) GLAXO GROUP LTD.
XX

PI
XX
Kimmerly WJ;

DR N-PSDB; AAH52588.

XX "nucleocapsid encoding polypeptides from *Staphylococcus epidermidis*
XX useful for vaccinating against infections, e.g. endocarditis -
PS Claim 18, Page 189, 2188pp, English.

AAH523040 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of the activity and therefore identify compounds that may be used for the treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed, S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH53971 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention. AAH55091 to N.B. The present invention specifically relates to the following:

no sequences are present for SEQ ID NO:4455 to 4464.

```
Query Match      5.2%      Score 140.5; DB 22; length 715;
Best Local Similarity 23.8%
Matches 105; Conservative 51; Mismatch 19;
```

26 TPVDELLANPAGEIMSTGYCQNDENRHSPLTQIT-----TENYGGQLVMARGM- 76
||| ||| : | : | :
309 tpct----kplfgekgeseptektkypvdeivhygseeikpqdhdfdomnncse 35

```

77 --PGK-----VOYPLIHDSGWYLANPQVYQAIDAKTGDILWEHRRQD--FNIA-- 122
      |||      ||| : : | || : : : : : : : : : : : : : : : : : :
365 dvppkpgvknptdgetvnttppvdvtekyppgvdpb1-----tsteelpfdckkrefdnlpag 420

```

[illegible]

```

      :: | | : ||| : | -----ASLVGKIPGGELGTWTRF 229
456 -----ekytkqpvdeivhyggseelkpgkhdefidpnarkysgtitpqky-----409
230 AVPDDGLIWRHQTLPDNDDECTFEPMVTV---INVDQSTNPGT 230

```

```

500 vknpdgtgvv-----tppvdd-----vkypvpgdplrtseelpfkckrefnpdl 545
282 ATGERRV-LTGP-CKT-GTMMQDFAETGEIARDJYNNIESTDENGIVTVEPADAI 318

```

Db 546 kpgervqkgeptkltitptknpitgkvgsepteiktkqvde--iteygsgeik 603
 QY 339 KELDEYDYVCPFLGGRWPS--AALNPDGS-IYFIPLANVCY-----DMAVNOE--F 387
 Db 604 pghkdefdpnarksgedvpgkpgvknptdgtvrvpvdvltkypvdqplstceelpf 663
 QY 388 TSMQVNTSNVTKLPGRKMT 408
 Db 664 dkrefnpd---lkpgkerv 680

RESULT 12

AAG83007 standard; Protein: 910 AA.

AC AAG83007;
 DT 03-SEP-2001 (first entry)
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:3108.
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KM vaccination; endocarditis.
 OS Staphylococcus epidermidis.
 XX WO200134809-A2.
 XX 17-MAY-2001.
 XX 09-NOV-2000; 2000MO-US30782.
 XX 09-NOV-1999; 99US-0164258.
 XX (GLAX) GLAXO GROUP LTD.
 XX Kimmery MJ;
 XX WPI: 2001-316495/33.
 XX N-PSDB: AAH53857.
 DR Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 PS Claim 18; Page 819; 2188pp; English.

PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 910 AA;

Query Match 5.28; Score 140; DB 22; Length 910;
 Best Local Similarity 25.48; Pred. No. 0.003; Indels 94; Gaps 23;
 Matches 98; Conservative 45; Mismatches 149;

QY 26 TPVTELLANPAPGEMISTGONENYRHSPLFOIT--ENV-----GQLQLVWANGMO- 76
 Db 231 tptl-----knpltgkvgsepteiktkqvdeiteygsgeikpghkdefdpnarksgse 286
 QY 77 --PGK-----VOVTEPLIDHGVWYLANPDVYQIADAKTGOLLMEHRRQL-PNIA-- 122
 Db 287 dvpqkpgvknptdgtvrvpvdvltkypvdqplstceelpf 342
 QY 123 TLNIVANGVIAGSTCOYSPFGCFVSGHDSATGEELMKNYFIPRAGEGDETWGNDYEAR 102
 Db 343 tekvvqg--epqtkltitp-----ltknplgkvy-----geqpt----- 377
 QY 183 WMTGAMGOIYTDVPTNLVHGSTAVGPA--SETQSTPGCTLXGNTPRFAVR-PDTGEIV 239
 Db 378 -----ekvtkqvdeivhygsgeikpghkdefdpnarksgedvpgkpgvknptdgtvrv 431
 QY 240 WRHOTLPDNDMDQECTEMWYTN--VDVOSTEMEGL-----QSTINPANG-ERYVLT 290
 Db 432 -----tpvdd-----vtkypvdqplstceelpfdkkrefdpnlapgtkvqk 477
 QY 291 GVP-GKT-GTMMQFPAETGEFLMARDTNYQMLIESIDENGIYVNDALIKELDEYDVC 348
 Db 478 geytkltitptknpitgkvgsepteiktkqvde--ivhygsgeikpghkdefdpn 535
 QY 349 PTFUGRDWPS--AALNPDGIFYIP 372
 Db 536 apksgedvpgkpgvknptdgtvrvp 561

RESULT 13

AAG82914 standard; Protein: 696 AA.

AC AAG82914;
 DT 03-SEP-2001 (first entry)
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KM vaccination; endocarditis.
 OS Staphylococcus epidermidis.
 XX WO200134809-A2.
 XX 17-MAY-2001.
 XX 09-NOV-2000; 2000MO-US30782.
 XX 09-NOV-1999; 99US-0164258.
 XX (GLAX) GLAXO GROUP LTD.
 XX Kimmery MJ;
 XX WPI: 2001-316495/33.
 XX N-PSDB: AAH53764.

DR Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 PS Claim 18; Page 763-764; 2188pp; English.

PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.

Fri May 24 11:27:24 2002

walick-934-125.pcp.rag

Job time: 381 sec

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001: 2001WO-US08631.

XX 31-MAR-2000: 2000US-0540217.

PR 23-AUG-2000: 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB: AAS90096.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 20: SEQ ID No 56268; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC to restore normal activity of (II) or to treat disease states involving

CC (II) (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG3037 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 948 AA:

XX

XX

XX

XX

XX

XX

Search completed: May 24, 2002, 10:19:18

Fri May 24 11:27:24 2002

walick-934-125.pep.rag

Page 12

Fri May 24 11:27:25 2002

wallick-934-125.pcp.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:03 ; Search time 39.53 seconds
(without alignments)
310.186 Million cell updates/sec

Title: WALICK-934-125.PCP
2689
Sequence: 1 MKPTSLMMSAGALALIAAP.....GMOYVATAGGSVYSGSLNS 502

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep.*
4: /cgn2_6/prodata/2/1aa/5B.COMB.pep.*
5: /cgn2_6/prodata/2/1aa/5A.COMB.pep.*
6: /cgn2_6/prodata/2/1aa/5B.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	17.7	754	4	US-09-296-284-4
2	474.5	17.6	738	1	US-07-985-458-3
3	467	17.4	720	4	US-09-296-284-25
4	312.5	11.6	739	3	US-09-136-251-2
5	105	3.9	1012	1	US-08-219-262B-10
6	105	3.9	1012	1	US-09-031-655-10
7	102.5	3.8	1721	3	US-08-700-651-5
8	102	3.8	1042	3	US-08-928-361B-11
9	102	3.8	1837	3	US-08-928-361B-5
10	99	3.7	3567	2	US-07-642-734C-4
11	99	3.7	3567	2	US-08-439-009A-4
12	98.5	3.7	500	2	US-08-987-519-2
13	98.5	3.7	966	2	US-08-868-786-2
14	96.5	3.6	635	4	US-08-931-608A-5
15	96.5	3.6	1043	4	US-08-928-361B-6
16	96.5	3.6	1721	3	US-08-868-786-6
17	96	3.6	974	2	US-08-868-786-6
18	95.5	3.6	548	5	PCT-US93-10541-2
19	95.5	3.6	583	4	US-09-401-476-2
20	95	3.5	532	3	US-08-737-336-6
21	95	3.5	606	3	US-08-883-534-3
22	95	3.5	606	3	US-09-204-764-3
23	95	3.5	714	2	US-08-472-534-3
24	94.5	3.5	1012	1	US-08-216-276A-19
25	94.5	3.5	500	2	US-08-987-519-1
26	94.5	3.5	512	4	US-09-356-818A-2
27	94.5	3.5	1112	2	US-08-714-402-2

28	94.5	3.5	1612	1	US-08-169-927-2	Sequence 2, Appli
29	93.5	3.5	348	1	US-08-247-902A-2	Sequence 2, Appli
30	93.5	3.5	816	1	US-07-731-157A-4	Sequence 4, Appli
31	93.5	3.5	816	1	US-08-229-444B-2	Sequence 2, Appli
32	93.5	3.5	816	2	US-08-541-780-4	Sequence 4, Appli
33	93	3.5	380	3	US-08-971-782-4	Sequence 4, Appli
34	93	3.5	380	4	US-09-309-026-4	Sequence 2, Appli
35	93	3.5	459	3	US-08-971-782-2	Sequence 2, Appli
36	93	3.5	459	4	US-09-309-026-2	Sequence 16, Appli
37	93	3.5	1381	4	US-09-540-245A-16	Sequence 8, Appli
38	92.5	3.4	824	4	US-09-626-589-3	Sequence 10, Appli
39	91.5	3.4	774	4	US-08-462-484-10	Sequence 10, Appli
40	91	3.4	527	1	US-08-441-147-10	Sequence 10, Appli
41	91	3.4	527	5	PCT-US95-07536-10	Sequence 10, Appli
42	91	3.4	551	2	US-08-793-229-32	Sequence 32, Appli
43	91	3.4	551	3	US-09-285-957-32	Sequence 2, Appli
44	91	3.4	651	4	US-09-626-589-2	
45	91	3.4	651	4	US-09-626-589-2	

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Lee, Eun-Hae
APPLICANT: Rhee, Sang-Ki
TITLE OF INVENTION: Glucosyltransferase subunits of Sorbitol dehydrogenase, Genes
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Glucosyltransferase subunits
US-09-296-284-4

Query Match	17.7%	Score 476	DB 4	Length 754
Best Local Similarity	26.3%	Pred. No. 2.1e-37		
Matches 158	Conservative	80	Mismatches 230	Indels 132
				Gaps 18
QY	14	LALLAAPFAQVTV	---	TDLLANPPAGEWISYGNONENRHSPLQITTEVNGQ 66
DB	17	LGCAAAALAFCAATSPVALAEDTGTATINADHDPDMNSYGRYSQYSPIDQITDKNSN 76		
QY	67	LOIYVARGMORGKQV	---	TPLIHGVMTLANPGVIOALDAKTCGLIWEHRQLP-NIA-- 122
DB	77	LKLAHYDLDINRQEGEPILVDGVMTATNWSKMAKDATKLMSTDPRVPGNADR 136		
QY	123	---TLNIVA---NGVYAGS---		---TCQYS 141
DB	137	GCDDIVNRGAAYWNGKVFEGTGDRLIALDAKTCGLIWSVYTPKBAQLGHRSTVDCA 196		
QY	142	-----FGC-----FVSGHDSANGEEIMKRYFIPRAGEED-----		----- 172
DB	197	PRIAKKVILIGGAGAEFGARGFVAYDAETGKMDMREFVTPNDKPDGAASDVLMSKA 256		
QY	173	---ETWMDYKRMWTG---AMGOITVDPYTNLVHGSTAVGPASFTQRTGPGCLYGTNR 228		
DB	257	YPTWVGAGMKGQGGGGVWDSILIDPYTDLYLGVGNSPMWKKFRSEKGNLFLGSI 316		
QY	229	FAVRPDTGELVWRHQTLPNDNDQCTEEMANTVNDQPSFEMGLOSINPNATGE-RR 287		
DB	317	VALNPDTGKTVMHFQETPMDDMDYTSVQOIMALDMPV-----NGEMRH 359		

```

QY 288 VLVGPKCTGTMMQFPAETGEFLMARDTNYOMIESIDE-NGIYTVEDALIKELDEYD 346
Db 360 VLVHAD-KNGFFYIIDAKTKRISKPTTYENWANGDPTGRNYPDGLMTLNGKPMY 418
QY 347 VCTPFGGRDMPASALNPDSCGYTFPLNNVCY-----DKMAVDOEFTSMDEV 393
Db 419 GIPGDLGGINFAMAYISQTKIVYIPAOQVPEFVVDPOKGFKAHDSMNLGLDMNKIGL 478
QY 394 NTSN----VTKLPGKMDIGRIDALDIDSTGRTLSVERAANSPVLSTGGVLENGCTD 449
Db 479 DUNDPOHAKADKQFLKDKGWIYAMDPOKQAAFTVDHKGPMNGGLATAGVLEFOGLAN 538
QY 450 RFRALSOETGELTMOFLATVASGAQISTEVDGMOYVAL---AG-----GGYSYSG 499
Db 539 GFHAYDTATGKDLFTPEPASAIITAPVYITANGKOYVAVEVGMSGITYPFLGVAKTSG 598

```

RESULT 2

US-07-985-458-3
Sequence 3, Application US/07985458
Patent No. 5344777

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takemura, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing the Same and Transformed Acetic Acid
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Fishauf, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268

INFORMATION FOR SEQ. ID NO. 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about

```

IDENTIFICATION METHOD: 72,000
ORIGINAL SOURCE:
ORGANISM: Acetobacter alioacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Takemura, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horiouchi, Sueharu and
AUTHORS: Heppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

```

Query Match 17.6%; Score 474.5; DB 1; Length 738;
Best Local Similarity 25.6%; Pred. No. 2,8e-37;
Matches 151; Conservative 77; Mismatches 229; Indels 133; Gaps 15;

```

QY 10 SAGALALLAPAFAYOTPTDE-----LIAPPAGEWISYQONENYHSPLOTIT 61
Db 17 TAGTICALSISATYASADDDGATGENAITHADHPGNMWTYRTYSDDQYSPLOINR 76
QY 62 ENYGOQLVWAKGMOQKQV-TPLIHGVWYLANPGVIOAIDAKGDLMEHRQLP- 119
Db 77 SNVGNLKLAMYLDLDTNRGOEGTPIVDGVYATTNSMKAVAAATGKLMSIDRVP 136
QY 120 NIA-----TLN----- 125
Db 137 NIDKCCDVTNRGAAYNNKGYFTGPDGLIALDAKTGLVSWVTIPPEAEIKQRSY 196
QY 126 ----YVANGYVAGSTCOYSPGCG--FVSGHDSATGEELMNRNYLPAGEGCD- 172
Db 197 TVDGAERIAKGRVILGN--CGSEFGARGFVSAPDAETGKDMFFVPPKNEPDAAS 254
QY 173 -----ETWGDYEAARMNGA-----NGQITDPTNLVHYGSTAVPASETOR 215
Db 255 VLMNKAYOTWS-----PTGAWTROGGGTWDSIVYDPAVLVYLGNGSPNNYKR 307
QY 216 GIPGTLGTYNTRAVRBDTGEIYWRHQLPRDMDQECTFEFMYTNYDVOPSTEMEGLQ 275
Db 308 SEGKGDNLFLGSIYALKRPERGEYVWHQETPMDQMDFTSDQOIMTLPLI----- 357
QY 276 SINPNAATGERRVLTGVPCKTGTMMQFPAETGEFLMARDTNYOMIESID-ENGIVTNE 334
Db 358 -----NGETRHYIYHARKNGFFYIIDAKTGEFISGRNYVYVWNASGLDPKTRPIYN 410
QY 335 DALIKELDEYDVCPTFLGGRDMPASALNPDSCGYTFPLNNVCYDMAVADDEFT-SNDVY 393
Db 411 DALYTLTGKEMVIGIPDGLGHNFAMAFSPKGLVYIIPAOQVPEFVLYNOYGGFTPHDSW 470
QY 394 NTSN----NVTKLPG-----KDMIGRIDALDIDSTGRTLSVERAANSPVLSTGGV 442
Db 471 NLGLDMNKVGIQDPEAKQAFYKDKGWIYAMDPOKQAAEMRVYDHKGPMNGGLATAGDGL 530
QY 443 LFNGTIDRYFRALSOETGELTMOFLATVASGAQISTEVDGMOYVALIAG 492
Db 531 LFQGLANGEFHAYDATNGSDFHFAADSGIITAPVYITANGKOYVAVEVG 580

```

RESULT 3
US-09-296-284-25
Sequence 25, Application US/09296284A

```

Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533.0870000
CURRENT FILING DATE: 1999-04-22
CURRENT APPLICATION NUMBER: US/09/296,284A
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
US-09-296-284-25

```

```

Query Match      17.4% Score 467; DB 4; Length 720;
Best Local Similarity 26.2% Pred. No. 1.4e-36;
Matches 152; Conservative 81; Mismatches 220; Indels 128; Gaps 18;

```

```

26 TPVTDELLAPPAEWISYSGOENYRHSPLDTITENVGOLVYVARGMOPGVYV-TP 84
5 TAITN---ADQHPDWMYSGRFYSQRYSPDLQITKDNNSKLKLAHYDLTNRQEGPT 61
85 LIHGVMTLANPGVIOAIDAKTGDLMEHRRLP-NIA-----TLNIVA--NGVIVA 134
62 LIVGVMTATNWSKRALDAATGKLMSTDPKVPGLADRGCCDVIYNGAAVWNGKYV 121
135 GS-----TCQYSP-----FCG 145
122 GFEDRLALDAKTKGLVWSYVTPKBAOLGHRSTYVDGAPRIAKGVYIGNGAEFGA 181
146 --FVSGHDSATGEELMRKYFIPRAGEED-----ETWGNDEYEAAMTG--A 187
182 RGFVATYADGKMKMREFYVNPDKPDGASDDVILMSKAYPTWKGAGMKDGGGQIV 241
188 WGOITPDTNVLVHGSTAVGPASSETORGTPGCTIGTNTFFAARPDGELVWHRHQLPR 247
242 WBSLIIDPTDLVYLGVGNGSPWNYKFRSEKGNLFLGSIVAINPDIGKIVHFOETPM 301
248 DMDQECTEMAVTNDVQPSSTEMEGLOSINPNATGE-REVLGVPCKTGTMMQDAET 306
302 DQMDYTSVOOIALDMPV-----NGEMRHVLVHAP-KNGEFYIIDAKT 343
307 GEFLMARDTYONMIESIDE-NGIVTNEDELKELDVEYDVCPTFLGGRMPSALNPD 365
344 GKISGKPYIYEWANGLDPVTRPNYNDALMTLNGKRWYGIPEGDLGSHNAAAYSPQ 403
366 SGFIYPLNNVCY-----DMAVDOEFTSMQVYNTSN---VTKLPPGKDMI 408
404 TKLYITRPAQVPPYVDPOKGGKFAHDSWNLGLDMNKIGLLDDNDPQHRADRAOKLKLK 463
409 GRIDAIISTGRILMSYERAAANYSPVLSTGGVLENGGTDRTFRALSDETELLMOTRL 468
464 GWIYAMPQKQAFTYVDHKGPMNGGILLATAGVGLFQGLANGEFHAYDATTKQLFTFPA 523
469 ATVASGQAISEYVDMQYVAI---AG-----GGVSYSG 499
524 QSATIAIPVITYANGKQIVAVENGWGSITPFLPLGAVARISG 564

```

```

RESULT 4
US-09-136-251-2
Sequence 2, Application US/09136251A
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: OJIMA, Setsuko
APPLICANT: SHINJOH, Masako

```

```

APPLICANT: TOMIYAMA, No. 612171561bumi
TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
CURRENT APPLICATION NUMBER: US/09/136,251A
CURRENT FILING DATE: 1998-08-19
EARLIER APPLICATION NUMBER: EP 97114432.4
EARLIER FILING DATE: 1997-08-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 739
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(24)
US-09-136-251-2

```

```

Query Match      11.6% Score 312.5; DB 3; Length 739;
Best Local Similarity 21.8% Pred. No. 1.6e-21;
Matches 149; Conservative 74; Mismatches 193; Indels 269; Gaps 33;

```

```

23 AOVTPVTDELLAPPAEWISYSGOENYRHSPLDTITENVGOLVYVARGM--OPG 78
86 SVPAMAPQSANPARGDWVAVYGRDDHQTYSPLSEITPENMSKLKLAFAVYHTGSIYPRG 145
79 KV-----QVTP-L-IHGVMTLANPGVIOAIDAKTGDLMEHR-----ROLPNIAF--L 124
146 QVNRKMAETTPPIKVGDLTYCSANDITK-LDPATGKQIMRRVNDVKYHSIPYTAACKGV 204
125 NIVANGVYVAGSTQYSPFCGFSGH-----DSATGEELMRNYFIRAG-- 169
205 TYFTSSVYBEGQPCN-----NRLIGTIDMLRLVAVDAETGD-----FCPWFGGGOVNL 254
170 -----EGDET-----GNDYEARMWTGAM-- 188
255 QGLGSEVGFVSMAPPVINGVVVNHVEYLDGGRAPASGVIRGYDAESGKFWAMADV 314
189 -----GOITYDPTNVLVHGSTAVGPASSETORGTPGCTIGTNTF-----FAVRPD 234
315 NSGRSOPAY-RVTYITVABRRIPGLPDRROR-----GSRLPDRNSAADYISALRSD 366
235 -----TGEIYVHRHQLPRDNDQECTEFEMAVTNDVQPSSTEMEGLOSINP 279
367 AERKVSAAVAIDVKTGSERWVFQAHKDVMDYDIGSATL-----MDMPG-----P 413
280 NAATGERRVLTGPCKTGTMMQFDETGELWARDNTYONMIESIDEN-----GIV----- 330
414 DGQTVPALIM-----PTKRGQTFVLDRTGK-----PLPVEERPADSPGVIPGDP 459
331 -----TYNEDAI--LKELD-----VEY 345
460 RSPQPSVGMPLARVPDLAKETDMMGMSPIDOLFCKIKRRANYGEFTPPSVDKPWLEY 519
346 DVCTPFLGGRDPSAALNPDGSIYFTPLN-NVCYDMAVDOEFTSM----- 390
520 ---PGYNGSDMSMSYDPSGILLANMNITPMYDOLVTRKADSLGLMPIDDPNPKRG 576
391 -----DYNTSNVTKLPPCKDMIGRIDAIDISTG--RTIMS--- 424
577 GGAEGANGMDGTPYGIIVTPEFMDY--TGMMCNRP-----YGMTALIDMHGKVLQHPDL 631
425 -----VERAANYSPVLSTGGVLENG-GTDRTFRALSDETELLMOTR 467
632 GTARANGPWGLPTGLPWEIGTPNNGSGSVYTGGLIFGATDQIRAIDEHTGKVVMSAV 691
468 LATVASGQAISEYVDMQYVAIAG 492
692 LPGGQAMPMTYEANGHOYVAIAG 716

```

RESULT 5


```

Db 114 PGVYALNGTINMT-----FQGSISEL-----TDYSYNGIMSATANIN 152
QY 325 DENGIVNEDAIKEIDVEYDVCPTFLGGRWPSAALNPDGSIYFIFILNNVC----- 377
Db 153 DKIGVNLVGBEVYVLSPTSYSLSYRLGD-PIPAAGLDP-----KLMTCDSDRPR 204
QY 378 -YDMAVADQ-EF-----TSMQDYNTSNYTKLPPGKMDI-GRIDAIDISTGRITLM- 423
Db 205 VYTVTADEYQFSSQLIPSGVKTFLFTANIDALTSLSVGELIFFSQVTHISEVDVTTIE 264
QY 424 -----SVERAANYSPVLSTGGVLFN--GGTDRTFALSOEGETLMQRLATVAS 473
Db 265 IGFDEGVYVKAATDFGLTGTNNLVFNLGPTSEITQPTISMLLEVYTYRGGT--A 322
QY 474 GOAISYEVDGMOYVAIAG 492
Db 323 GDPISWTSGLATVTVGG 341

```

```

RESULT 7
US-08-700-651-5
; Sequence 5, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUN, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-4(HV) US/08/700,651B
; CURRENT FILING DATE: 1997-08-14
; EARLIER FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1721
; TYPE: PRN
; ORGANISM: Cryptosporidium parvum
; US-08-700-651-5

```

```

Query Match 3.8%; Score 102.5; DB 3; Length 1721;
Best Local Similarity 20.5%; Pred. No. 1.6; Indels 231; Gaps 37;
Matches 135; Conservative 78; Mismatches 216;

```

```

QY 25 VTPYDELLANPPAGEWISYGNQ-----NYRHSPLTQITTEVNGOLOLWARKQOP 77
Db 979 IDPTTG-LPBNPPTGLHINFTNNMTDSSPAGAYKXAAVANGIKTDNV-----YGLPV 1029
QY 78 GKVOYVPLIHGVMYLANPGDVIOAIDAKTGLIMEHRQOLNIAVTLNVA----- 128
Db 1030 GEITGLP-----KDPGSDI-PFNSTGELVADPSTGKPIINNSAGVSKPGAPLPIE 1079
QY 129 --NGVI-----VAGSTQYSP-FGCFVSGHDSAT-----GEEIMRNFIIP--RAGEE 170
Db 1080 DENGINLFDESTNLPIDGNNQVLNPEINSTYSGSTGTTKPKPIPVNGGAYVDEFAKQD 1139
QY 171 GDETMGNDYEAARMWTGANGQITYPDVYNLHYGSTA--VGPASPTORGFG----- 219
Db 1140 ADK--GKD--GLIVPPTNSIKRDKPVNTQYSNITGININP--ETGKIVIGSLPGSLNYP 1192
QY 220 -----GLVGTFRFAVRPDTGELIWRHQLPRDN--WQECTFEWMTN-- 262
Db 1193 SFNTPQOTDEITIKPVDTYTGLPYDSTGEIIDPATKLPIDGSAVAGEILTEVINITIDE 1252
QY 263 ----VDVQ--PSTEMEGLQS-----TNP-----NAATG--ERRVLT 290
Db 1253 VTGLPIDLETGLPRDPVSGILPOLPNGTIVDPNSKKPIGSGHSGFINGTSGEOSHEDKDPST 1312

```

```

QY 291 GVPCKTGT-XMOFP-----AETGEFLMARDNTY-----QNMIES 323
Db 1313 GKPLDPTNGLHPFDEDSGLINLPETGDKLOGSHSGFMPVPGKPOGENGIMTPQOILEA 1372
QY 324 IDENGIVTVNE-----DAI-----LKEIDVEYVCTFLIG----- 354
Db 1373 L--NKLPISNENINISPRSDAVDPDRPTNTWNNKISQOTYQVDSKTTILSASVIFHAL 1430
QY 355 ----RDWPSAAL--NPDGSIYFIFLNNVCYDMAVADQ-----EFTSMQDYNTSNYTKLPP 403
Db 1431 GPTQPTDPTTGLPDSPTGLPPIPGFNVLYDPQCEQIKGSPVYSLYVEKKNITVEAY 1490
QY 404 GKMICRIDAIDISTGRITLMSVERAANYSPVLSTG-----GGVLFNGGTDRTYFRA-- 454
Db 1491 GL-----PVDPKGTGFPIDPI-----STLPPAKNGELIDPISKYSISGSIAGISKAG 1538
QY 455 ----LSOETGETLMQF-----RLATVASGOAISYEVDGMOYVAIAG-----GGV 494
Db 1539 SOSKSDSEGNPIDPSTNMFYDPKGGKILDPESGIAIDNSYSGV-FATVPTAARKGGV 1597

```

```

RESULT 8
US-08-928-361B-11
; Sequence 11, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESS: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1042 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-361B-11

```

```

Query Match 3.8%; Score 102; DB 3; Length 1042;
Best Local Similarity 20.9%; Pred. No. 0.79; Indels 232; Gaps 36;
Matches 138; Conservative 71; Mismatches 219;

```

```

QY 25 VTPVTEDELLANPAGEMISYGNONE-----NYRHSPLTITTEENVGOLQVWARGMOP 77
DB 301 IDPTTG-LPFNPPTGHLINPNNNTMSSFGAKYKIVASNGIKTDVNYGLPVDIEITGLPK 359
QY 78 GKQVTPPLIHGQVMTLANPCVIOAIDAKTGDLIMEH-----RQLPNI-----121
DB 360 DPVSDIP-----FNSITGEL-----VDPSTGKPIINNNTAGIVSGKGLPPIEDENGFLD 409
QY 122 --ATLNIYANGVI-----VAGST--COYSPGCFVSGHDSATGEELMRYEIPRAG 168
DB 410 PSTKLPIDGNQVLPNPTNSTVSGSTGSKRPK-GLPVNCGGVVDEE-----AK 459
QY 169 EEGDETMGNDYEARMMGAMQITTDYVTLVHYGSTA--VGPASETORGTG-----219
DB 460 DQADK-GKD--GLIYPPNINSIKDPVNTQYSNTGNIINP--ETGKVIPLGSLPGLN 512
QY 220 -----GTLYGNTFRFAVRPDGTGEIWMRHQTLPRDN--WDQCTEEMAVTN 262
DB 513 YPSENTPOQDEITGKRPVDTVGLPYDPSTGELIDPATKLPISGVADDELITVLENTT 572
QY 263 -----VDVQ--PSTMEGLQS-----INP-----NAATG-----ERRV 288
DB 573 DEVTGLPIDLEITGLPRDPVSGLPQLPNGTLVDPNSKKRPIPSHSGFINGTSGQSHKDP 632
QY 289 LTGVPECKTGMQFD-----AETGEFLWARDNTYQNM--ESIDENGIYVNEADAIL 338
DB 633 STGKPLDPNTGLPFDEDSGSLINPETGDKLGSGHSGTFMVPKPGCEGNGIMTPBQ-IL 691
QY 339 KELDV-----EYDVC-----TFL--GGRDWPSSAL-----362
DB 692 EALNKLPTSEVNISPRSSDAVDRPTNTWMNKISQOTFOVQKKTIPGSAASYHTAL 751
QY 363 -----NPDGTYIFIPLANVCYDMAVDO-----EFTSMDYNTNSNTKLP 403
DB 752 GTPPTQDPTTGLPSPDPTGFLPIGFEVNLVDPQTGQIGSVPIYSLYKEKNITVEAY 811
QY 404 GKDMIGRIDAIDISTGRTIWSVERAANYSPLVSTG-----GVLFNGGTDYRFA---454
DB 812 GL-----PVDPKTGPPIPI-----STLPRAKNGELIDPISGYFSGSIAGFISGAG 859
QY 455 -----LSOETGETLMOT-----RLATVASGOAISYEVGMOYVIALG-----GV 494
DB 860 SQSSSDSGNPIDPSTNMPYDPKTKLIDPESGIALDINSVSGV-FATVPGTAAPKKGCV 918

RESULT 9
US-08-928-361B-5
; Sequence 5, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRKA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: CA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, HANA
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1837 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-5

```

```

Query Match
Best Local Similarity 3.8%; Score 102; DB 3; Length 1837;
Matches 138; Conservative 71; Mismatches 219; Indels 232; Gaps 36;

```

```

QY 25 VTPVTEDELLANPAGEMISYGNONE-----NYRHSPLTITTEENVGOLQVWARGMOP 77
DB 1096 IDPTTG-LPFNPPTGHLINPNNNTMSSFGAKYKIVASNGIKTDVNYGLPVDIEITGLPK 1154
QY 78 GKQVTPPLIHGQVMTLANPCVIOAIDAKTGDLIMEH-----RQLPNI-----121
DB 1155 DPVSDIP-----FNSITGEL-----VDPSTGKPIINNNTAGIVSGKGLPPIEDENGFLD 1204
QY 122 --ATLNIYANGVI-----VAGST--COYSPGCFVSGHDSATGEELMRYEIPRAG 168
DB 1205 PSTKLPIDGNQVLPNPTNSTVSGSTGSKRPK-GLPVNCGGVVDEE-----AK 1254
QY 169 EEGDETMGNDYEARMMGAMQITTDYVTLVHYGSTA--VGPASETORGTG-----219
DB 1255 DQADK-GKD--GLIYPPNINSIKDPVNTQYSNTGNIINP--ETGKVIPLGSLPGLN 1307
QY 220 -----GTLYGNTFRFAVRPDGTGEIWMRHQTLPRDN--WDQCTEEMAVTN 262
DB 1308 YPSENTPOQDEITGKRPVDTVGLPYDPSTGELIDPATKLPISGVADDELITVLENTT 1367
QY 263 -----VDVQ--PSTMEGLQS-----INP-----NAATG-----ERRV 288
DB 1368 DEVTGLPIDLEITGLPRDPVSGLPQLPNGTLVDPNSKKRPIPSHSGFINGTSGQSHKDP 1427
QY 289 LTGVPECKTGMQFD-----AETGEFLWARDNTYQNM--ESIDENGIYVNEADAIL 338
DB 1428 STGKPLDPNTGLPFDEDSGSLINPETGDKLGSGHSGTFMVPKPGCEGNGIMTPBQ-IL 1486
QY 339 KELDV-----EYDVC-----TFL--GGRDWPSSAL-----362
DB 1487 EALNKLPTSEVNISPRSSDAVDRPTNTWMNKISQOTFOVQKKTIPGSAASYHTAL 1546
QY 363 -----NPDGTYIFIPLANVCYDMAVDO-----EFTSMDYNTNSNTKLP 403
DB 1547 GTPPTQDPTTGLPSPDPTGFLPIGFEVNLVDPQTGQIGSVPIYSLYKEKNITVEAY 1606
QY 404 GKDMIGRIDAIDISTGRTIWSVERAANYSPLVSTG-----GVLFNGGTDYRFA---454
DB 1607 GL-----PVDPKTGPPIPI-----STLPRAKNGELIDPISGYFSGSIAGFISGAG 1654
QY 455 -----LSOETGETLMOT-----RLATVASGOAISYEVGMOYVIALG-----GV 494
DB 1655 SQSSSDSGNPIDPSTNMPYDPKTKLIDPESGIALDINSVSGV-FATVPGTAAPKKGCV 1713

RESULT 10
US-07-642-734C-4
; Sequence 4, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:

```


Fri May 24 11:27:25 2002

walick-934-125.pcp.ra1

Page 7

APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

Query Match 3.7%; Score 99; DB 2; Length 3567;
Best local similarity 21.9%; Pred. No. 12;
Matches 122; Conservative 49; Mismatches 211; Indels 174; Gaps 25;
QY 21 ARAQVPTVDELNPAGW-----ISYQNGENRHSPLQITTEVNGQLQVW 71
DB 2911 ARAPIAVTDHRLARVPG--WSDADAAVPIAY--TTAHVALHDLGLRAGOSVLIH 2964
QY 72 ARGMPGKVQVYTLIHGVMYLANPG----- 97
DB 2965 AAGGVMAAVALARRAGAEVLTATAGPAKGTLRALGLDDEHIASSRETFARKFRERTG 3024
QY 98 ----DVIQAIIDAKTGLIWEHRQLPNTATLNIYANGVIV-AGST-----COYSFF 143
DB 3025 GRGVYV--LNSLTGLIDES-----ADL-LAEDGVFVEKGTDLRDAGDFRGYAPF 3074
QY 144 GCFVSGHSAATGELMWRNYFIPRAGEGDETWGNDYFARMNTGAW-----GOITYDPVTLN 199
DB 3075 DLGKAG--DRLGELLIRVYVGLAGELD-----RLPVSAWELGSAPRALQHSRSG 3123
QY 200 VHYGSTAV-GRASETORGT---PGCTLVGTNRFAVAPPTGEIYWRQTL-PRDNMDQEC 254
DB 3124 RHVGLVLTGPAVPDPDGTGLITGGT--GTGLRLARHLVTEHGVRRHLLVSRGAGAPG 3181
QY 255 TTEWMTVAVDQSTEMEGLOSINPNAAT---GERRVLTGVPCKTGTMMQFDETEGFL 310
DB 3182 SELETAIEDLGAETIACOTADRALSLALDGLRPLTGVAAGVUL----- 3230
QY 311 WARDNTYQNTIESIDENQVIT-VNEDAILKELDEYDVCPTELGKRWPSAALNDSGIT 369
DB 3231 ----ADGLVTSIDEPRAVEQVLRKAVDAA-----VNLHELTANTGLS 3267
QY 370 FIFLNNVCTDMAVDOEFTSMYVNTSNV-----TKLPPGK----- 405

DB 3268 FEVLSSASAVIA-----GPGGVYAAANESLNAALARRRGLPAKALGMLRAGASEMT 3323
QY 406 ----DMIGRIDIDISTGTLMSVERAANYSPVLSTGGVLFNGTDRYFRALSOETGE 461
DB 3324 SGLGDRIAR-----TGVAALPTEFRLALFDSALRGRGEVVEPLSINRSALRRAEVPE 3376
QY 462 TL-WOTRLATVYASGOA 476
DB 3377 VLKGMVRAKLRAAGOA 3392

RESULT 11
US-08-439-009A-4
Sequence 4, Application US/08439009A
Patent No. 6004787
GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
TITLE OF INVENTION: Specific Polypeptides
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952.US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-439-009A-4

Query Match 3.7%; Score 99; DB 3; Length 3567;
Best local similarity 21.9%; Pred. No. 12;
Matches 122; Conservative 49; Mismatches 211; Indels 174; Gaps 25;
QY 21 ARAQVPTVDELNPAGW-----ISYQNGENRHSPLQITTEVNGQLQVW 71
DB 2911 ARAPIAVTDHRLARVPG--WSDADAAVPIAY--TTAHVALHDLGLRAGOSVLIH 2964
QY 72 ARGMPGKVQVYTLIHGVMYLANPG----- 97
DB 2965 AAGGVMAAVALARRAGAEVLTATAGPAKGTLRALGLDDEHIASSRETFARKFRERTG 3024
QY 98 ----DVIQAIIDAKTGLIWEHRQLPNTATLNIYANGVIV-AGST-----COYSFF 143
DB 3025 GRGVYV--LNSLTGLIDES-----ADL-LAEDGVFVEKGTDLRDAGDFRGYAPF 3074
QY 144 GCFVSGHSAATGELMWRNYFIPRAGEGDETWGNDYFARMNTGAW-----GOITYDPVTLN 199
DB 3075 DLGKAG--DRLGELLIRVYVGLAGELD-----RLPVSAWELGSAPRALQHSRSG 3123

```

QY 200 VHYGSTAV-GPASETORG-----PGTLGNTNTPFAVPDGTGLVWRHQT-L-PRDMWDOEC 254
DB 3124 RHVGKLVLPAPVPDPDGTGLVITGT--GTLGRLLARHLVTEGVHLLVSRKADPG 3181
QY 255 TFEWMTVNDVOPSTEMEGLOSINPNAT-----GERRVLTGVPCKGTGYMOPDAETGFL 310
DB 3182 SPELAETEDGASAEIACDTPADRLSALLDGLPRPLTGVHAGVL----- 3230
QY 311 WARDNTYONMIESIDENGIVT-VNEAAILKEIDVEYDVCPTFLGGRDPSAALNPDGTY 369
DB 3221 -----ADGLVTSIDEPVAVQVLRKVDA-----WNHLELANTGLS 3267
QY 370 FIPINNYCYDMANVDEFTSMYYNTSNV-----TKLPCK 405
DB 3268 FFLVJSSASVLA-----GRQGVYAANESLINALALRTGDLPAKALGMLAQAASEMT 3323
QY 406 -----DMIGRIDALIDISTGTLMSVRAAANYSPVLSTGGVLENGGTDYFRALSOETGE 461
DB 3324 SGLGDRIR-----TGVAALPTERRALALPDSALRGRGEVFPFLINSALRRAEFYPE 3376
QY 462 TL-WOTRLATVASGA 476
DB 3377 VLNGWVRAKLRAGA 3392

```

RESULT 12

```

US-08-987-519-2
; Sequence 2, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Luderer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987.519
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRP
; ORGANISM: Human Papillomavirus
US-08-987-519-2

```

Query Match
Best Local Similarity 3.7%; Score 98.5; DB 2; Length 500;
Matches 56; Conservative 30; Mismatches 88; Indels 89; Gaps 12;

```

QY 36 PRAGEMISYGONOENRHSPLTQITTEENVGQLOLVWARGMOPGKYQVPLPHDGVMLAN 95
DB 158 PPLGHWKKGKQCTN--TP-----VOAGDCCPLELITSVIADGDM----- 196
QY 96 PGDIYQALDAKTGDLIWEHRROLPNATLNIYANGV--IVAGSTCY-----SPFG- 144
DB 197 -----VDGFGAM-----NEADLOTKSDVPIDICGTCKPDYLOMAADPYGD 240
QY 145 -----CEVSGHDSATGELMRYEIPRAGEBDETMGNDYEARMWMTGANGQITTYDPVNLVH 201
DB 241 RLFFFL-----RKQEMFARHFFNRAGEVGPV-----PDLIIRK 274
QY 202 YGSTAAGPASEFROGTGTLTGTNTPFAVPDGTGLVWRHQTLPDQ--WDQECTEEM 258
DB 275 GSGNRTSVASSTLYVTPTGSLVSEADLFNKP-----YWLQKQGHNNICWGN-----QL 325
QY 259 MVTNVDVOPSTEMEGLOSINPNA 281
DB 326 FVTVDVTRSTNMTLCASVTTSS 348

```

RESULT 13

US-08-868-786-2
; Sequence 2, Application US/08868786

```

; Patent No. 5998701
; GENERAL INFORMATION:
; APPLICANT: Kawchuk, Lawrence M.
; APPLICANT: Armstrong, John
; APPLICANT: Lynch, Dermot
; APPLICANT: Knowles, Richard
; TITLE OF INVENTION: Potatoes Having Improved Quality
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5730 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80803
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 04-JUN-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,946
; FILING DATE: 10-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Sally A
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 8-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-868-786-2

```

Query Match
Best Local Similarity 3.7%; Score 98.5; DB 2; Length 966;
Matches 84; Conservative 56; Mismatches 157; Indels 153; Gaps 18;

```

QY 25 VTPYTDLANPRAGEMISYGONOENRHSPLTQITTEENVGQLOLV-----WARGM 75
DB 613 VTPRMTIRFCNPPLSAITTKTGTEDM-----VLKTEKLAELQKPADNEDLQNEWREAK 666
QY 76 QPKGVQVTPPLHDSGVWYLANRGDIYQALDAKTGDLIWEHRROLPNATLNIYANGVIYAG 135
DB 667 RSNKIKIVSEFLKEKTYISVP-DAMFDIOYK--IHEYKROLNT-----FG 710
QY 136 STCYQSPCEVSGHDSATGELMRYEIPRAGEBDETMGNDYE--RMTGAMGQI 191
DB 711 IYIRTKM-----KEMTAERKTN-FYPRVCIFGKAFATYVQAKRIVFITDVGATI 762
QY 192 TYD-----PYTNLVHGSTAVGPASETOR--GTFG-----G 220
DB 763 NHDPEIGDLKVPVPPYNSVAELLIPASDISEHISTAGMEASGTSMMKFAAMGCIQIG 822
QY 221 TLTYGNTNTPFAVPDGTGLVWRHQTLPDQWDOECTEFEMVYTNVDVOPSTEMEGLOSINPN 280
DB 823 TLDGANVE--IREVEG----- 837
QY 281 AATGERVLTGVPCKGTGMOPDAETGFLNARDNTYONMIESIDENGIVTVEAAILKE 340
DB 838 -----ENFFLEGAHBIAGLRKERADGKFV--PDREFEVKEFVRSGARGSYNYDOLIS 891
QY 341 LD-----VEIDVCPPTLGGHWPASALNPDGTYFIPLNVCYDMANVDEFTSMOY 393

```



```

QY 25 VTPVTDELLANPAGEMISYQONE-----NYRHSPLTQITTEBNVQLOLVNARCMOP 77
Db 301 IDPTTG-LPFNPTGHLINPTNNMTDSSFAKAYAVSNGIKTDNV-----YGLPV 351
QY 78 GKVOYTPLIHGVMTLANPGDVIAIDAKTDGLIWEHRROLPTATLNIVA-----128
Db 352 GETTGLP-----KDPGSDI-PFNSTGTGLVDPSTGKPINNSTAGIVSGKGLPPIE 401
QY 129 --NGVI-----VAGSTCOYSP-FGCFVSGHDSAT-----GHEIMRNYFIP--RAGEE 170
Db 402 DENGTLFDPSTNLPIDGNQOLVNPETNSIVSGSTGTTKPKPGIPVNGGCVVDEEAKDQ 461
QY 171 GDETMGNDYEXARMWTGAMGOITTPVTLNLYGSTA--VGPASETORGPFG-----219
Db 462 ADK--GKD--GLIVPPTNSINKDPVTNTQYSMTGNIIINP--ETGKVIPLGSLPGSLNP 514
QY 220 -----GLYGTNTRFAVRPDTEIYVRHOTLPRDN--WDOCTEFEMAVTN--262
Db 515 SFNTPQOTDELTKGKPVDTVGLPYDSTGELIDPATKRLPGSVAGDEILTEVLNITTDE 574
QY 263 ----VDVQ--PSTENEGLOS-----INP-----NAATG-----ERRVLT 290
Db 575 VTGLPIIDLETGLPRDPVSGILPOLPNCGLTVDPsNKKRIPGSHSGFINGISGEOSHEKDPST 634
QY 291 GVPCKTGT--MMOPD-----AETGEFLMADTNY-----ONMIES 323
Db 635 GKPLDNTGIAHPFDEDESGSLINPETGDKLQSGSHSGTFMPVPGKPOGENGIMTPEQILEA 694
QY 324 IDENGIVTYNE-----DAI-----LKELDVEYD-----VCPTEL 352
Db 695 L--NKLPTSNENVISPRSSDAVDPDPTNTWMMNKISGQTYQVDGKKTIPGSASAVIHIAL 752
QY 353 G--GRDMPsAAL--NPDGSIYFIPLANNYCYDMAVDO-----EFTSMDVYNTSNVTKLPP 403
Db 753 GTPOTDPTTGLPSPSTGLPIPGFNVLVDQGEQIKGSVPYSLYVEKENYVTEAAY 812
QY 404 GKDMIGRIDAIDISTGRILMSVERAANYSPVLSTG-----GGVLFNNGGTDRYFRA---454
Db 813 GL-----PVDPKTGFPIDPI-----SYLPFAKNGELIDPISGKIYSGSIAGFISGKAG 860
QY 455 ----LSOETGETLMOT-----RLATVASGOAISYEVDGMQYVAIAG-----GGV 494
Db 861 SOSKSSDESGNPIDSTNMPYDPKGLIDPESGIAIDNSVSGV-FATVPGTAAPKKGAV 919

```

Search completed: May 24, 2002, 10:20:08
 Job time: 321 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:21:03 ; Search time 54.76 Seconds
(without alignments)
880.877 Million cell updates/sec

Title: WALICK-934-125.PEP
Perfect score: 2689
Sequence: 1 MKPTSLIMWASAGALLALAAAP.....GMQYVAINGGVSYGSGINS 502

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	18.1	742	2 A49340	alcohol dehydrogenase
2	481	17.9	601	2 B95863	alcohol dehydrogenase
3	479.5	17.8	708	2 S52317	alcohol dehydrogenase
4	475	17.7	742	2 J50326	alcohol dehydrogenase
5	474.5	17.6	738	2 S14270	alcohol dehydrogenase
6	435.5	16.2	626	2 J00706	alcohol dehydrogenase
7	413	15.4	623	2 B83399	alcohol dehydrogenase
8	391.5	14.6	573	2 S68591	alcohol dehydrogenase
9	339.5	12.6	796	2 JY0107	alcohol dehydrogenase
10	339.5	12.6	796	2 H90644	alcohol dehydrogenase
11	339.5	12.6	796	2 H85495	alcohol dehydrogenase
12	331.5	12.3	796	2 AG0523	alcohol dehydrogenase
13	315	11.7	801	1 S00943	alcohol dehydrogenase
14	287.5	10.7	808	1 OPEX	alcohol dehydrogenase
15	271	10.1	809	2 B98314	alcohol dehydrogenase
16	271	10.1	809	2 A12968	alcohol dehydrogenase
17	267	9.9	803	2 F83360	alcohol dehydrogenase
18	258	9.6	778	2 G98221	alcohol dehydrogenase
19	258	9.6	778	2 A13064	alcohol dehydrogenase
20	250.5	9.3	809	2 A55547	alcohol dehydrogenase
21	230	8.6	639	2 J04881	alcohol dehydrogenase
22	209.5	7.8	524	2 A41378	alcohol dehydrogenase
23	153.5	5.2	407	2 A82580	alcohol dehydrogenase
24	140.5	5.1	668	2 C75264	alcohol dehydrogenase
25	137.5	5.0	839	2 D97013	alcohol dehydrogenase
26	137	4.9	827	2 F64512	alcohol dehydrogenase
27	134.5	4.9	866	2 A82284	alcohol dehydrogenase
28	132.5	4.9	392	2 F91050	alcohol dehydrogenase
29	130.5	4.9	392	2 F91050	alcohol dehydrogenase

30	130.5	4.9	392	2 B85895	probable dehydrogenase
31	130.5	4.9	392	2 AG0821	probable dehydrogenase
32	129.5	4.8	392	2 G65027	probable dehydrogenase
33	126.5	4.7	525	2 A89953	probable dehydrogenase
34	125.5	4.7	593	2 S00128	probable dehydrogenase
35	125	4.6	407	2 C82804	probable dehydrogenase
36	125	4.6	2535	2 AC0304	probable dehydrogenase
37	124	4.6	3972	2 S75251	probable dehydrogenase
38	121.5	4.5	943	2 J04081	probable dehydrogenase
39	120	4.5	1354	2 AG0538	probable dehydrogenase
40	118.5	4.4	799	2 T48889	probable dehydrogenase
41	116.5	4.3	424	2 T29127	probable dehydrogenase
42	116.5	4.3	1350	2 AF2005	probable dehydrogenase
43	116	4.3	1156	2 C87371	probable dehydrogenase
44	116	4.3	1441	2 B86807	probable dehydrogenase
45	115.5	4.3	1197	2 D82696	probable dehydrogenase

ALIGNMENTS

RESULT 1
A49340 alcohol dehydrogenase (EC 1.1.1.1) precursor - Acetobacter pasteurianus (strain NC113)
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteurianus
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:DJ13893; NID:9517067; PIDN:BAA40252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyridoxal-5-phosphate-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match	18.1%	Score 487	DB 2	Length 742
Best Local Similarity	26.5%	Pred. No. 3.5e-27		
Matches 161	Conservative 81	Mismatches 215	Indels 150	Gaps 21
QY	11	AGALALLAAPAFVQVPTVDL--ANPAGMISYGNQNGEYRHSPLTQITTEENVGQD	68	
DB	23	AAALPYAAYPARADGGGTGEALIHADHPENMLSTGRTYSQRSPLDQINRSVGLK	82	
QY	69	LVARGMQPKVQ--VPLINDGYVLANPGDYIALDAKTGLIWEHRQLP-NIA----	122	
DB	83	LAWYTTLDITNGOEAITPLVVDGIMVATTNMSKKEALDAATGKILMOYDKPGNTADKGC	142	
QY	123	--TLN-----	126	
DB	143	CDTVNRGAGYWNKGVFNGTDFDGLVAADAKTGKWEVNTIPADSLGKQSYVDGAVR	202	
QY	127	VANGYIVAGSTQYQSPFC--FVSGHDSATGEELRMNYFIPRAGEGD-----	172	
DB	203	VAKGVLVGN--GSEFARGFSAFPADEGKIKMRYVTPNNKNRBPDAVADNVLMSKA	260	
QY	173	-ETWGNDEYARMT-----GAMGQITDPTNTNIVHSTAVGPASFTQRTPGTGLY----	223	
DB	261	YKTWGP--KGAVRGGGGGTWDSLVYDPSDLI--LAVG-----NGSPNNYKRS	309	
QY	224	--GTN-----TRAVRPDGEIWRHQTLPRNMDOCTEFEMATNDVQSTMEGLQSI	277	
DB	310	GIGSNLEFSGIYALKRPEIGEVYVHFOATPMQMDVTSVOQITMLDMV-----	357	
QY	278	NPAATGE--RRLVGVPCKTGTMOQDAETGEFLMARPTNQNNTIESIDE--NGIYVNE	335	
DB	358	-----NGEMRHVIMHAP--KNGEFYVLDATGTEFLGKRYVQNNANGIDPLTGRPIYND	411	
QY	336	ALLKEIDVEIDVCPFLIGRDWPSAALNPDGIVFIPLNNCY-----DMAA	382	

Query Match	17.98;	Score 481;	DB 2;	Length 601;
Best Local Similarity	25.84;	Pred. No. 6.9e-27;		
Matches 155;	Conservative 90;	Mismatches 206;	Indels 150;	Gaps 22
QY	14 LALLAAPAFAYVTVDDEL--LANPAGMISYSGQNOENYRHSPLQITTEENGGOLWV 71			
Db	8 LAIMSIGGAQVAFANDELQKLIIDP-NOMALQTDGVDANLEYSKIDQINRDNVKGQLQVAM 66			
QY	72 --ANGMPGVQVYTPLIHGGVMTLANP-----GDV----- 99			
Db	67 TFSFGVLRGH-EGSPVLIDMLVYHPFNVTYALDLSKQIQIYWKPEPKQDNPVIVMC 125			
QY	100 -----IQAIAKQIGDLIW-----EHRROLPNATENTIVANGV 131			
Db	126 CDPIYRCGVAYADNKKIFLHQADITTYVALDAADKAGKIYASVKNKGDATKGETNATVATMPVKQKI 185			
QY	132 IVAGSTCOQYSPFGCFVSGHDSATGEELRNFY-----PRAGEEGE 173			
Db	186 LVGISGGEFGVG-HYTAYSMADGKVLMKSGVSMKPSDILLIDPEKTHLGRKPVGDQSLT 244			
QY	174 TWKNDYEAARMMTG---ANGQITDYPVTNLIYHSGTAVGAPASEQRQRTPGGLIYGINTFPA 230			
Db	245 TWEGD---QMKIGSGGTGWGVSYPDENLIYVYGTGNPSTWPNFR--PEDNRW-SMTIFA 298			
QY	231 VRDTEIYWRKQITPLRPNMDQCTEFEMATVNDVDPSTEMEGLOSINPNATGRRRLT 290			

```

Query Match          17.8% ; Score 479.5; DB 2; Length 708;
Best Local Similarity 25.0% ; Pred. No. 1,1e-26;
Matches 148; Conservative 84; Mismatches 228; Indels 121; Gaps 18;

QY      3  PTLSLMASGALLMLAPFAQVTPVT-----DELLANP-AGEMISYGNQEN 50
      1  :::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      12  PGRWVLLAACLG-SAAAFQGTGPAQAQAAAVQVDSDFRANAATPMPWPTGVAYE 69
      1  :::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY      51  YRHSPLQIITTEWVGQLOLVMARGQPK-VQVPLIHGVMYLANGVQIAIDAKTGD 109
      1  :::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      70  TRYSRLDILANAAVNDKLGAMSYNLESTRGEATPVVVDGIMVYSKSYVAHIDRTGN 129
      1  :::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY      110  LIWEHRROLPNI-----ATLNIYANG----- 130
      1  :::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      130  RIMTYPQIDIRSTGFGCCDDVNRGVALMKRYVYGAMDCRLALDAAGKEWQHNTFE 189
      1  :::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```


alcohol) dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - *Methylobacterium*
N:Alternate names: methanol dehydrogenase 62k large chain
C:Species: *Methylobacterium extorquens*
C:Date: 31-Mar-1992 #sequence=revision 31-Mar-1992 #text=change 08-Oct-1999
R:Accession: J00706; S07908
R:Anderson, J.M.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
Gene 90, N/73-16, 1990

A: Reference number: J00706
A: Accession: J00706
A: MUID: 90337342

A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA

A/residues: 1-626 <AND>
A/Cross-references: GB:M31108; NTD:q150017; PIDN:AAA25380.1; PTD:q150018
A/Experimental sources: EMBL:U00096

R; Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989

A1>Title: The second subunit of methanol dehydrogenase of *Methylobacterium extorquens* AM
A1Reference number: S04644; M01D:89350892

A;Molecule type: protein
A;Residues: 28-50 'xy' 5

Comment: This enzyme oxidizes methanol to formaldehyde.

Genetics:
Gene: MOXF
Keywords:

<SIG>
#status predicted
#signal sequence
Domain: signal sequence
Product: alcohol dehydrogenase

 acceptor | dehydrogenase | (acceptor) | #status experimental | <MAT> |

Query Match
Best Total

best local similarity 24.68; pred. No. 1.4e-33;
Matches 148; Conservative 91; Mismatches 208;
Indels 155. Gaps 33.

10 SAGALALLA-AP-AFAQVTPVDELLANPAGE-WISYGONENRHSPLQOTTENGO 66

```

67 7 SVSALMLALAPALSSGAYANDKLVELSKSDDNWVMGKNYDSNNFSDLKQINKGNVQ 66

```

7 100 -----

126 RAVACCDLVNRLAYWPGDGKTPALIKTQLDGNVAALNAETGETVWK-----

```

0Y 126 IVANGIVASTCOYSPF-----GC-----FVSGHDSATGPELNRNFI----- 164
Db | : | | : | :
174 -VENSIDIKVSTLTLLIAPYVKDVIIGSSGAEILGVRGLTAVDVKTGEQVRRATANGPDK 232
0Y 165 -----PRAGEE--GDEWGNDYEANRMTGA---WGQITDPVNLVHYGSTAV 207
Db | : | | : | | | : | | | | : | :
233 DLLIASDFNIKNHYQOKGLGTGTWEGD---AMKIGGGTNGWYAYDPGNTLLYFGUNP 289
0Y 208 GPASEIQRGTGGTLYGNTPREAVRPDGEIYWRHOTLRPNMDQETFEEMATINDVOP 267
Db | : | | : | | | : | | | : | | | : | :
290 APNNEIEMR--PGIDNKW-TWITIGRADADGEAKFGYQKTPHDEM-----YAGVNMW 338
0Y 268 STEMEGLOGINNATGERVLTGVPCKGTGMOFDAETGEFLMAR---DTYQNNIESI 324
Db | : | : | : | : | : | : | : | : | :
339 LBSQOK-----KDGARKLTLHPDRNGIYVLTDRTDGALVSAKKIDT--VAVFKSV 388
0Y 325 DENGIVYVEDALIKLEID-VEYDVCPTFLGGRPMSAALNPDGSIFFIDPLNNVCYDMNAV 383
Db 389 DUKTQOPVRDEPYGT RMDHLAKDLCPSAMKYNHOGHDSYDPKRELFPMGINICHMDWEPF 448
0Y 384 DQEFMSMIVYNTSVTKLPQK-----DNIGRIDAIIDISTGRTLSVSRBAANSPYLS 437
Db 449 MLPIYAGCFEFGALNNYPPGKGRDQNYEELQIKAYNMITDIDYEMKEKERRAAYWGCTMA 508
0Y 438 TGGGVFNGGTDRYFRALISOETGETLMQRLNLTAVASGALISEVDGMOYVAIAGGVSYG 497
Db 509 TAGDLVEFGTLDTIKARSDSDTDLMKKRIPSGALIGYPMYTFHKGTQYAI-----YVG 563
0Y 498 SG 499
Db |
564 VG 565

```

```

b03399
quinoprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain ATCC 27803)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: B83399
C/Storer: C

```

adman, S.; Yan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardy, K.; Lory, S.; Olson, M.V.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.;

Nature 406, 959-964, 2000

A:reference number: A82950; PMID:20437337
A:accession: B63399
A:Status: preliminary
A:Title: *Detonifera* PAU1, an opportunistic pa-

A: Molecule type: DNA
A: Residues: 1-623 <STO>

A: Cross-references: GB:AE004624; GB:AE004091; MID:99947973; PIDN:AAG05370.1; GSPDS:GN
A: Experimental source: strain PA01

C/Gene: exaA; PA1982
 A/Gene: exaA; PA1982

Query Match

Best Local Similarity	12.78;	score 43;	DB 2;	length 623;
Matches 140; Conservative	23.3%;	Pred. No. 6e-22;	Mismatches 245.	T=3.1e-100

	Indels	Gaps
3 PTSLINASAGALAL--LAAPAFQVTPVDELLAN--DPAGEWISYGNQNFYRSDPT 57	240	122

9 P A G L L R P S L H C I A F A V A L G S A G A A I A K D V T W E D I A N D K T T G D V L Q Y G M G T H A Q R W S P L K 68

```
Y      58 QTTTENVGQLQVLWARGM---QPCKVQVTPLIHDCVMYLANPGDVIQAIDAKTGDLWE 113
```

```

09 QVNADNNEKLFAMWSYSGDEKQRGQ-ESQAIVSDGVIIY7ASYSRFLALDADTKGRLLWT 127
114 HRPOTDN-----

```

128 YNRLPPDIDRCDDVYNRGAATYGDKEVECTTDACTCTT.....IATL-NIVA-NGVIV--AGST 137

138 CQYSD-----EGCF--VSGHDSATGEFLBNYETDP-----
139

Db 268 WGYKTPHDEMDAGVQAVMLTD-----QPVN-----GKMTPLSHIDRNGIL 330
QY 300 WQDAETGEELMADTN-YQNMIESIDENGIVTVNEDAILKELDVE-YDVCPPTFGGRDW 357
Db 331 YTLRENGNLLVAKRVPAPVAVKFKVDLKTGTVPBDEFATRMDRHGTNCPSPAMCFHQ 390
QY 358 PSALNPDSGYGTFPLNNVCYDMAVDOEFTSMDVYTSVYTLP-----PGKMIGRIDA 413
Db 391 GVDSDTPSRLLYGLNGLICMDMEFMLPRAGOFVGAITLAMPGNGFTKEMQIRIRA 450
QY 414 IDISTGRTLSVERAANYSPVLSGGGVLENGGTDTRFRALSQETGETIMQTRLATVAS 473
Db 451 PFLTGTGAKWTKWEKFAAMGGTLTYTGGVYATLLDGLKLDKDKCKELMNEFMPSGGI 510
QY 474 GQALISEVDGMQYVALAGGVSYSGSG 499
Db 511 GSPMTYSFKGQIT-----GSMKGVG 531

RESULT 9
JV0107
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Escherichia coli*
C/Species: *Escherichia coli*
C/Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C/Accession: D64735, JV0107, A45997, S45201: I41228
R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.

glucose dehydrogenase (pyrroloquinoline-gulonone) (EC 1.1.99.17) - *Escherichia coli*
JVO107
C:Species: *Escherichia coli*
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64735, JVO107; A45997; S45201; I41228
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
..; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <BIAT>
A:Cross-references: GB:AE000122; GB:U00096; NID:91786315; PIDN:AACT3235.1; PID:917863
A:Experimental source: Strain K-12, substrain MG1655
R:Clifton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding *Escherichia coli* quinolone
A:Reference number: JVO107; MUID:91035240
A:Accession: JVO107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51323; NID:941553; PIDN:CAAS3706.1; PID:941554
A:Experimental source: strain K12
R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Purification and characterization of aminoproteinase from *Escherichia coli* O157:H7

A:Reference number: A45997
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <I>AM>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45161
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <I>U>
A:Cross-references: EMBL:D26562; NID:9473770; PIDN:BA05580.1; PID:9473791
A:Experimental source: strain K-12 substrain W3110
R:Yamada, M.; Asakura, S.; Sailer, M.H.; Yamada, Y.
J. Bacteriol. 175, 568-571, 1993
J:Title: Characterization of the *gcd* gene from *Escherichia coli* K-12 W3110 and regula
A:Reference number: I41228
A:Accession: I41228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'KRRCRT', 157-193, 'K', 194-665, 'H', 667-796 <I>RS>
A:Cross-references: GB:D12651; NID:9216555; PIDN:BA02174.1; PID:9216556
C:Genetics:
A:Gene: *gcd*

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline-quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM1>

F:63-81/Domain: transmembrane #status predicted <TM2>

F:96-110/Domain: transmembrane #status predicted <TM3>

F:120-140/Domain: transmembrane #status predicted <TM4>

F:93/95/Binding site: ubiquinone #status predicted <TM5>

F:466/Active site: Asp #status predicted

Query Match
Best Local Similarity 12.6%; Score 339.5; DB 1; Length 796;
Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;

```

QY 6 LMAS-----AGALLAARFAQVTPYDELANPAGEMWISYCONENRHSPLTQ 58
DB 135 LTMAGFNDPQELNGTSLADTPAEA-ISPVADQ-----DWPAYGRNDEGQFSPKQ 185
QY 59 ITTENVGOLQVMA-----RGMOPKV--QVTPLEHIDGVMTLANPGDVIQAIADAKTD 109
DB 186 INADNVHNKEMAVFRTGDKQPNPDGEITNEVPIKVGDTLYLCTAHQRLFALDASGR 245
QY 110 LIMEHRRL-----
DB 246 EKMHYDPELKTNEFQHTCGVSYHEAKETASPEVMADCPRIILPVNDRRLAINAE 305
QY 119 -----PNTATLIYAN-----GVVAGS--TCQYS--PGCFV 147
DB 306 NKLCEETANKGVNLQSNMPTKRGLEYEPTSPITDKTIYVAGSVTNESTRETSGVI 365
QY 148 SGHSATGEELMRYNFIPIRA-----GEGDETKGNDYEARMGACQIYDPTVLY 200
DB 366 RGFVNTGELLMA--FDPGAKDPNAIPSDHTFTNS-----PNSWAPAAYDAKDLV 416
QY 201 HYGSTAVGPASETORPGTGLTGTMTTRAVRPDGEIYWRHQLPRNMDOECTFEEMV 260
DB 417 YLPKGVTPPTDVGKNTPEDEYASSI-LALNATGKLMASYQVYHDLMD-----468
QY 261 TNDVOSTEMEGLOSINPNAATGERRVLTGVCKTGTMQDAETGEL-----310
DB 469 --LEAQT-----LADITYN--GQKVPYIAPAKTGIFVLDNRNGELVYPADEKVPQ 518
QY 311 WARDNYONMIESIDENGI-----VYVEDAILKEI--DVEYD-----346
DB 519 GAAKGDVTPPTQFSELSRPTKDLGADMWGATMFQDLYCRVAFHQRREGIFTPPSQ 578
QY 347 ---VCPTEFLGGRDWSAALNP-----D 365
DB 579 GTLVFPGNLMGEMGIVDPNREVAIANPVALPFVSKLIPRGGNMEQPKDAKTGTE 638
QY 366 SGI-----YFPLNNCYDMAVAVDEFTSMVYNTSKLPKDKMIGRIDAIDIST 418
DB 639 SGIOQOYGVPIGVTLN-----PFLS-----PFGLPCKQAPMGYISALDLKT 679
QY 419 GRTLSVERAANYS-----PV-----LSTGGVLPFGCGT-DYFPALSOE 458
DB 680 NEVWKRRIGTPDQSMFPMPVPVPEFMGMPLGPISTAGNVLFIAATADYTLRAYNMS 739
QY 459 TGETLMOTRLATVASGA--ISYVDGMQYVAI--AGGSVSYGS 498
DB 740 NGEKIMOGRLP--AGGATPTMYEVNGKQYVVISAGHSFSGT 780

```

RESULT 10
H90644
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: H90644
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
A:reference number: A59629; MUID:21156231; PMID:11258796
A:Accession: H90644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA33551.1; PID:g13359584; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: Ecs0128
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match
Best Local Similarity 12.6%; Score 339.5; DB 2; Length 796;
Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;

```

QY 6 LMAS-----AGALLAARFAQVTPYDELANPAGEMWISYCONENRHSPLTQ 58
DB 135 LTMAGFNDPQELNGTSLADTPAEA-ISPVADQ-----DWPAYGRNDEGQFSPKQ 185
QY 59 ITTENVGOLQVMA-----RGMOPKV--QVTPLEHIDGVMTLANPGDVIQAIADAKTD 109
DB 186 INADNVHNKEMAVFRTGDKQPNPDGEITNEVPIKVGDTLYLCTAHQRLFALDASGR 245
QY 110 LIMEHRRL-----
DB 246 EKMHYDPELKTNEFQHTCGVSYHEAKETASPEVMADCPRIILPVNDRRLAINAE 305
QY 119 -----PNTATLIYAN-----GVVAGS--TCQYS--PGCFV 147
DB 306 NKLCEETANKGVNLQSNMPTKRGLEYEPTSPITDKTIYVAGSVTNESTRETSGVI 365
QY 148 SGHSATGEELMRYNFIPIRA-----GEGDETKGNDYEARMGACQIYDPTVLY 200
DB 366 RGFVNTGELLMA--FDPGAKDPNAIPSDHTFTNS-----PNSWAPAAYDAKDLV 416
QY 201 HYGSTAVGPASETORPGTGLTGTMTTRAVRPDGEIYWRHQLPRNMDOECTFEEMV 260
DB 417 YLPKGVTPPTDVGKNTPEDEYASSI-LALNATGKLMASYQVYHDLMD-----468
QY 261 TNDVOSTEMEGLOSINPNAATGERRVLTGVCKTGTMQDAETGEL-----310
DB 469 --LEAQT-----LADITYN--GQKVPYIAPAKTGIFVLDNRNGELVYPADEKVPQ 518
QY 311 WARDNYONMIESIDENGI-----VYVEDAILKEI--DVEYD-----346
DB 519 GAAKGDVTPPTQFSELSRPTKDLGADMWGATMFQDLYCRVAFHQRREGIFTPPSQ 578
QY 347 ---VCPTEFLGGRDWSAALNP-----D 365
DB 579 GTLVFPGNLMGEMGIVDPNREVAIANPVALPFVSKLIPRGGNMEQPKDAKTGTE 638
QY 366 SGI-----YFPLNNCYDMAVAVDEFTSMVYNTSKLPKDKMIGRIDAIDIST 418
DB 639 SGIOQOYGVPIGVTLN-----PFLS-----PFGLPCKQAPMGYISALDLKT 679
QY 419 GRTLSVERAANYS-----PV-----LSTGGVLPFGCGT-DYFPALSOE 458
DB 680 NEVWKRRIGTPDQSMFPMPVPVPEFMGMPLGPISTAGNVLFIAATADYTLRAYNMS 739
QY 459 TGETLMOTRLATVASGA--ISYVDGMQYVAI--AGGSVSYGS 498
DB 740 NGEKIMOGRLP--AGGATPTMYEVNGKQYVVISAGHSFSGT 780

```

RESULT 11
H85495
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933

Db 753 GQATPMTYEVNGKQYVVISAGHGSFGT 780

RESULT 13

S00943

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoac

C:Species: Acinetobacter calcoaceticus

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S00943

R:Clifton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.

Nucleic Acids Res. 16, 6228, 1988

A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase

A:Reference number: S00943; MID:88289368

A:Accession: S00943

A:Molecule type: DNA

A:Residues: 1-801 <CLE>

A:Cross-references: EMBL:X07235; NID:g38711; PID:CAA30222.1; PID:g38712

A:Experimental source: strain LMD 79.41

C:Function: A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:9-35/Domain: transmembrane #status predicted <TM1>

F:39-57/Domain: transmembrane #status predicted <TM2>

F:61-79/Domain: transmembrane #status predicted <TM3>

F:94-108/Domain: transmembrane #status predicted <TM4>

F:118-137/Domain: transmembrane #status predicted <TM5>

F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted

F:471/Active site: Asp #status predicted

Query Match

Best Local Similarity 11.7%; Score 315; DB 1; Length 801;

Matches 148; Conservative 76; Mismatches 192; Indels 282; Gaps 30;

```

QY 20 PAFQVPTVTELLANPAGEMISYXGONENYHSPLOTTTNNVQOLQVARGMOPK 79
Db 152 PETAQAVGVAE-----SDMPAYGRTQAGVAKRISPKQINDQNVLDKVAWT--LRGD 202
QY 80 V-----QVPLHGVWYLANPGDVYQAIIDAKTGDLIWEHROU----- 118
Db 203 LKTNDGSETTNOVTPRIKINNMFTCAHQOLIAIDPATGKEMRRDPRKLDKSFQHLT 262
QY 119 ----- 118
Db 263 CRGVYTDANNTEFATSLQSSKSSQCRKVFVPVNDGLVAVNADTGACDFFGNG 322
QY 119 -----PNTATLNIYANGVIVAGS--TCQYS---PFGCVFSGHDSAT 154
Db 323 QVNLQEMFPAIYPGYNFTSPGIVT---GSTVVIAGSVTDNYSNKEPSG--VINGYVNT 377
QY 155 GEBLMRYEPIRAGE-----EGDETWGNDYBARMTGAMQOITYDPVTNLVHYGSTAVG 208
Db 378 GKLLM--VFDTGADPDNAPMAGEGTFVINS-----PAMAPLAVYDKKLIIV--VPTGV- 427
QY 209 PASETGRTGPGGILYG-----TTPRAVBPDEGIEIYWRHQTLPNNMOECGFEM 258
Db 428 -----GTP--DIMGDRETELKERYANSLAINASTGKLWNQOTHHMDLMDVPSQP 478
QY 259 MYTNNDVQDSTEMEGLOSINP--NAATGERFVLGVP-----CTKG--TMMQF 302
Db 479 SLADIRKNAGQTVPAIYVLTGTGNAFVLDK--NGQPIYVTEKRPVQGYKRPQTKGEF 536
QY 303 DAETGEF-----LMA-----RDTNYQNNIESIDENGITYVNE 334
Db 537 YSKTPQFSDMLAPQDLTKDQMLDQMLCQVSKRLNADGITYTPPSNGTL----- 592
QY 335 DAILELDEYDVCPTEFLGSDWPSALNPS-----GIYFT----- 371
Db 593 -----VPGNLGVFENGKMSVNPDRQVAVAMPDGLPFWSRLLPADPNRAQTAK 640
QY 372 -----PLNNCYMAAVALDOEFTSMQVYNTSNVTKLPKGDIMIGRIDAIIDISTGRTL 422

```

Db 641 GAGTEGQVDPMPGVY-----GVEISAFSLPGLPCQKOPAKRYVAGVDLKTHEV 690

QY 423 W-----SYERAANYSPLYLSTGGVYLFNFGCT--DRYFALSQETGEL 463

Db 691 WKRICTIRDSLPLNLFQUPAVKIGVPGIGSSISTRAGNMVGVATQDNYLRAFNVTNGKTL 750

QY 464 MQRRLATVASGA--ISTVEDQMOTVAL--AGGVSYGS 498

Db 751 WEARLP--AGGATPMTYEINGKQYVIMAGHGSFGT 786

RESULT 14

QPKEX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda

C:Species: Gluconobacter oxydans

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998

C:Accession: S17716; S19265

R:Clifton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.

Mol. Gen. Genet. 229, 206-212, 1991

A:Title: A single amino acid substitution changes the substrate specificity of quinop

A:Reference number: S17716; MID:92017653

A:Accession: S17716

A:Molecule type: DNA

A:Residues: 1-808 <CLE>

A:Cross-references: EMBL:X62710

R:Goosen, N. submitted to the EMBL Data Library, February 1992

A:Reference number: S19265

A:Accession: S19265

A:Molecule type: DNA

A:Residues: 1-212, 'A', 214-808 <GOO>

A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417

C:Genetics: A:Gene: qdh

C:Function: A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran

F:9-28/Domain: transmembrane #status predicted <TM1>

F:35-54/Domain: transmembrane #status predicted <TM2>

F:60-76/Domain: transmembrane #status predicted <TM3>

F:94-110/Domain: transmembrane #status predicted <TM4>

F:122-138/Domain: transmembrane #status predicted <TM5>

F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted

F:470/Active site: Asp #status predicted

Query Match 10.7%; Score 287.5; DB 1; Length 808;

Best Local Similarity 19.9%; Pred. No. 1e-12;

Matches 140; Conservative 82; Mismatches 229; Indels 251; Gaps 25;

```

QY 13 ALALAAPAPAAQVPTVTELLA-----NPPAGEMISYXGONENYHSPLOTITT 61
Db 128 AVLALFASLFTDPHDISGELPQIANASPADBDNVPASEMAHYGRTQAGDWSPLNDINA 187
QY 62 ENVGQLOLV-----ARGMOPKAV--QVPLHGVWYLANPGDVYQAIIDAKTGDLIN 112
Db 188 TIVSNLKVAMHHTTKMMANSNDPGQOTNEATPIEFNNLTLYWCSLHQKLFVADGATGVNKA 247
QY 113 EHRROU--PNTATLNI-----VANGYIV----- 133
Db 248 VYDPKLIQINPQHILCRGVSFHEPRANAMSDGNPAPTPDCANDSLIPVNDGLRLEVAD 307
QY 134 AGSTC-----QYSPFGCV----- 147
Db 308 TGTCTSGGNGNEIDLKRVNPQPTTPGQYEPTSPVITDKLIANSALDINSYKQASCA 367
QY 148 -SGHDSATGEELM-----FNYFIPRAGEGDETMGNYEARMTGAMQOITYDPVTNLVHY 202
Db 368 TQAFVYVTGKRWVWFDASNPDPNQLPDESHPVHPNPSMVIS-----SYDANILNLYI 422

```

```

OY 203 GSTAVGPASEFQKPT---GGTLAGTNTFR-----AVRPDTEGIWNRQOTLEPRDWMDEC 254
Db 423 PMGV-----GTPDWMGSDRTKDSERAPGVALNMDTKLAFYOTVHHNDIMDEL 473
OY 255 TPEMMVTVNDVOPSTEMEGIOSINPNNAATGERVLYGVPCKTQTMQOPAEFG-EFIMAR 313
Db 474 PQSPSLVDTQKDDTLVPAIYA-----PRTGDIIVLDRKKEIVPAP 517
OY 314 DT-----NYQNMISIDENGIVTVNEDAILKELDV-----EYD 346
Db 518 EEPVPOGAAPGHTSPPTQMSQ---LTLRPKNPLNDSIDMGRTIFDQECISYFHTLRE 574
OY 347 -----VCPFLGGRDWPFSALNPDGSIYFPLNNVCCDMAVYO----- 385
Db 575 GPTPLSLKSLIFPDGLQMFEMGGIADVPOQVAFANPISLPVSQULVRRGPNLWPE 634
OY 386 -----EFTSMDYNTSNVTKLPPGKDMI-----GRIDAIDISTGRTL 422
Db 635 ENNAKGTGGTGLQHNAYGYAVNLHPFLDPVLLPFGIKMPCRTPPWGVAGIDIKTKKV 694
OY 423 WSVERRAANYS-----PV-----LSTGGVYLF-NGGDRYFRALSOETG 460
Db 695 WQHNNGTLRDSMYGSSLPPLPPIKIGVPSLGGPLSTAGNLGFLTASMDYIIRAYNLTTG 754
OY 461 ETLWOTRLATVAVSQAISYEVDGMQVVAIAGGVSYSGSLNS 502
Db 755 KVLWODRLPAGCAATPTTVAINGKOYI-----VTYAGGHS 790

```

RESULT 15

```

B98314
Probable quinate dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: B98314
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tun
A:Reference number: A97359; PMID:11743194
A:Accession: B98314
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-809 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK90036.1; PID:g15160013; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_2936
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```

```

Query Match 10.1%; Score 271; DB 2; Length 809;
Best local similarity 21.6%; Pred. No. 1.6e-11;
Matches 152; Conservative 70; Mismatches 224; Indels 258; Gaps 30;

```

```

OY 8 WASAGATALLAAPAF-----AQTPTVDELLANPPAGMISYGQNGENY 51
Db 128 FAVAAVALAAGAGFAGFVAHPTVAFNGEPALTPVKPE-----TEQKNWEHYGNTSGGS 183
OY 52 RHPSPLOTTTENVQOLQVW-----ARGMQPKVQVYPLIHGVMYLANPDDVIOA 102
Db 184 RFVALDEITRDNINKLEAMTYHTGDPFISPANGAEDQOTPLQVGDVFLCTPHNNVIA 243
OY 103 IDAKTGDLIWE-----HR 115
Db 244 LDADGTETKWKAEIKSSKSSVMRCGLAYPDAKAPLKQTPAGSTPYTAANAEGALCOR 303
OY 116 ROLPNTAIALNIVA-----NG-----VIV 133
Db 304 RILMTINIELALDADGACPDPTGRTKGRDALKIGMDADDPQYVLTSAPTLAGTIVV 363
OY 134 AG--STCOYSPFGCFVSGHDSATGEELMRNYFTPRAGEEDFTW-----GNDYEARMM 184

```

```

Db 364 GERIADNVQVMPGVMRGFDVYTGELRW-----AFDEGNBDITKLPPEPGQTY-TRST 415
OY 185 TGAAGQITVDPTNLVH--YGSTAVGPASEFQKPTPGGLTYGTRFRAVRPDDGEIWMR 242
Db 416 PNWMAWSYDPDLNTVFMFVPGSPSVLYGATR--TPLDHYGA-SMLALDATTGREKMY 472
OY 243 QTLPRNNQOECEFFEMAVTVNDVOPSTEMEGIOSINPNNAATGERVLYGVPCKTQTMQOPAEFG-EFIMAR 313
Db 473 QTVHNDLMD-----FDVPMQPS-----FVDFPKADG--TSVPALVFGTKAG 511
OY 298 TMOQFAETG-----EFLMARDTNYQ-----MIESIDENGIVTVNEDAILKELDV-----EYD 346
Db 512 QLYVLDRTAQOPLTKVEEVSVKANIPNEPYLPTQPRSVGMQIAGOTLFTADMMGATPF 571
OY 335 DAILKEK--DVEYD-----VCPFLGGRDWPFSALNPDGSIYFPLNNVCCDMAVYO----- 385
Db 572 DOLLCRIAFKMRREGIYTAAGTDLAFLPFGSLGGMMWGSLSTDPYTTITVNDKRLGLW 631
OY 372 -----PL-----NNVCYDMAVDOEFTSMDYNTSNVTKL-----PRGKDMIGRI 411
Db 632 IEMKAAPTAKAVASGESVNTGMGVPMKGTPTAVYKNRFLSALGIPCQAPP-----YGTW 687
OY 412 DAIDISTGRTLSVERAANYS-----PV-----LSTGGVYLF-NGGDRYFRALSOETG 460
Db 688 TALDKKTOQIKMOPVYGVTEDTGPIGIMGLPPIGMPPLTGLTLAGTLAGGGLVFTAGTODY 747
OY 452 FRALSOETGETLWOTRLATVAVSQAISYE--VDGMQVVAIAGGG 493
Db 748 LRAFDTATGKEVMKARLPVGSQGGPMSTKSPKTKGQVIVASAG 791

```

```

Search completed: May 24, 2002, 10:21:06
Job time: 344 sec

```

Fri May 24 11:27:26 2002

wallick-934-125.pep.rpr

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:50 ; Search time 28.06 Seconds
(without alignments)
692.702 Million cell updates/sec

Title: WALICK-934-125.PCP
Sequence: 1 MKPTSLMWSAGALALLAAP.....GMOYVAINGGVSYGSLINS 502

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	17.8	739	1 DHET_ACEU	044002 acetobacter
2	475	17.7	742	1 DHET_ACEAC	P18278 acetobacter
3	474.5	17.6	738	1 DHET_ACEPO	P28036 acetobacter
4	461.5	17.2	757	1 DHET_GLUOX	O05342 glucobacter
5	435.5	16.2	626	1 DHM1_METEX	P16027 methylolact
6	435.5	15.9	600	1 DHM1_METOR	P15279 methylolact
7	426.5	15.8	631	1 XOXF_PAKDE	P29968 paracoccus
8	424	15.4	623	1 DHM1_PAKDE	P12293 paracoccus
9	413	14.4	571	1 EXAA_PSEAE	O9417 pseudomonas
10	388.5	14.4	766	1 DHM1_METME	P38539 methylolact
11	339.5	12.6	801	1 DHG_ACOXI	P15877 escherichia
12	315	11.7	801	1 DHG_ACOXI	P03465 acetobacter
13	287.5	10.7	808	1 DHG_ACOXI	P27175 glucobacter
14	250.5	9.3	808	1 DHG_ACOXI	O59086 acetobacter
15	240	8.9	790	1 DHG_ACOXI	O60782 acetobacter
16	134.5	5.0	827	1 YFGL_ECCLI	O62872 methanococc
17	129.5	4.8	593	1 YFGL_ECCLI	P27274 escherichia
18	123.5	4.7	799	1 AFSK_STRCP	P17974 streptomyc
19	118.5	4.4	1355	1 AFSK_STRCP	P54741 streptomyc
20	115.5	4.3	353	1 RPOD_ANASP	P22705 anabaena sp
21	115.5	4.3	353	1 RPOD_ANASP	O42111 bacillus su
22	115.5	4.3	353	1 RPOD_ANASP	O42111 bacillus su
23	115.5	4.3	353	1 RPOD_ANASP	O42111 bacillus su
24	114	4.2	796	1 Y232_RICPR	O92482 rickettsia
25	112	4.2	796	1 Y232_RICPR	O92482 rickettsia
26	110.5	4.1	3317	1 FLEY_CAUCR	P58365 rattus norv
27	108.5	4.0	3354	1 CADN_HUMAN	P15345 caulobacter
28	108.5	4.0	3354	1 CADN_HUMAN	O91251 homo sapien
29	107.5	4.0	1323	1 ASPA_AERSA	P28897 strawberry
30	106	3.9	621	1 ASPA_AERSA	P31339 aeromonas s
31	106	3.9	621	1 ASPA_AERSA	P07140 drosophila
32	104	3.9	697	1 NAME_STRPN	O54727 streptococc
33	104	3.9	697	1 NAME_STRPN	O54727 streptococc

34	103.5	3.8	595	1 YE85_SCHPO	O14301 schizosacch
35	103	3.8	1637	1 MRSP_STRAU	P80544 staphylococ
36	102	3.8	638	1 PAOX_ARTGO	P46881 arthrobacte
37	102	3.8	790	1 POLG_IBDVO	P06867 sus scrofa
38	101	3.8	1012	1 GTF5_STRDO	P27276 avian infec
39	101	3.8	1365	1 GTF5_STRDO	P29336 streptococ
40	100.5	3.7	448	1 SPGL_STRSP	P06654 streptococ
41	100.5	3.7	466	1 MM08_RAT	O88766 rattus norv
42	100	3.7	465	1 MM08_MOUSE	O70138 mus musculu
43	100	3.7	566	1 AMY_STRGR	P30270 streptomyc
44	100	3.7	1276	1 SCAP_CRIGR	P97260 cricetus
45	99.5	3.7	807	1 AFSK_STRGR	P54742 streptomyc

ALIGNMENTS

RESULT	ID	STANDARD	PRT	739 AA.
1	DHET_ACEU			
AC	044002: 007952;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).			
GN	ADH.			
OS	Acetobacter europaeus.			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;			
OC	Glucobacter.			
OX	NCBI_TaxID=33995;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-DES11 / DSM 6160;			
RA	Thurner C.A.K.;			
RL	Submitted (NCBI-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBUNIT: HETEROETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO			
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +			
CC	reduced acceptor.			
CC	-1- COFACTOR: PO4 AND HEME (BY SIMILARITY).			
CC	-1- SUBUNIT: HETEROETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO			
CC	SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE			
CC	COMPLEX (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC			
CC	SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib.ch).			
CC	-----			
CC	EMBL: X82894; CAA58066.1; -			
CC	DR EMBL: Y09480; CAA70688.1; -			
CC	DR HSSP: Q92477; IFLG.			
CC	DR InterPro: IPR001479; Bac_POO.			
CC	DR HSSP: Q92477; IFLG.			
CC	DR InterPro: IPR002372; Bac_POO_repeat.			
CC	DR InterPro: IPR000345; CytC_heme_bind.			
CC	DR Pfam: PF01011; Bacterial_POO_1; 1.			
CC	DR PROSITE: PS00364; BACTERIAL_POO_2; 1.			
CC	DR PROSITE: PS00190; CYTOCHROME_C; 1.			
CC	DR Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.			
CC	KW SIGNAL			
CC	FT CHAIN 1 35			
CC	FT CHAIN 36 739			
CC	FT BINDING 651 651			
CC	FT BINDING 654 654			
CC	FT BINDING 655 655			
CC	FT METAL 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;			
CC	FT SEQUENCE			

Query Match

17.8% Score 479; DB 1; Length 739;

Best Local Similarity 25.9%, Pred. No. 2.9e-28;
Matches 151; Conservative 77; Mismatches 236; Indels 120; Gaps 15;

```

OY 10 SAGALALLAARPAQVPTVDE-----LLANPAGEMISYGOENVRHSPLQIT 61
DB 17 TGTGICALISGYATMASADGOGATEAIHADDHGMNMTYGRISQORSPIDQINR 76
OY 62 ENVGLOLVARGMOPKQVYV-TPILHGVYLANGVYQAIADAKTGLIWEHRRLP- 119
DB 77 SNVGNLKLAWYLDIDTRNGEGTPIVDGMYATTWMMKAVDADATGLMSYDPRVG 136
OY 120 NIA-----TLN----- 125
DB 137 NIADKGGCDYVNRGAATWNGKYTFEGFDRLIALDAKTGLIWSVNTIIPRAELGKORSY 196
OY 126 -----IVANGVYVASTCOYSPFGC--FVSGHDSATGEBELMRNFETIRACEEDDETWGN 177
DB 197 TVDGAPRIAGRYIIGN--GSGFEGARGFVTFADAEITKVDMPREFETAPNPKNEPDHYSAD 254
OY 178 D-----YEARWMTGA-----WQITYPVTNLVHGSTAVGPASETGRTGRT 221
DB 255 SVLMNKAYOTWSPGTGAMTNGGQGVYWSIYDPAVDLVYGVNGSGPMNYKTRSEKGD 314
OY 222 LVGTNFRVAVRDTGEIYWRHQTLPDRNWDCECTEFEMVTVNDVQPTSEBGLQISINPA 281
DB 315 NIEFLSIALKPEETBEYWHQETPMQDFTSVQOIMTIDLPLNGET----- 362
OY 282 ANGERRVITGVPCKTGTMMQDFAEGEIMAROTNYONMIESID-ENGIVYVNEADILKE 340
DB 363 -----RHVIVHAP-KNGFEYIIDAKTGFIISGKNVYVNASGIDPRTGPIYNPALVTL 417
OY 341 LDVEYDVCPTFLGGRDPSAALNDSGIYFIPLNVCYDMMAVDQFET-SMUYVNTS--- 396
DB 418 TGKEMVYIGIPDLGHNRAAFSPKGTGLVYIPAOQVPLTYTNOVGFTPHDPSMVLGDM 477
OY 397 NVTKLPRG-----KDMIGRIDAIDISTGRTLSVEEAANYPVYSTGGVLENGCT 448
DB 478 NKVGIPSPPAKQAFYVDLKGKIVAMDPOKQAEARVYDHKGFWNGGILATGDLFOGIA 537
OY 449 DRYFRALSOETGELMOTRLATVASQAISYVDMQYVAAIAGG 492
DB 538 NGEFHAYDATNGSDLFHRAADSGIIPVYTLANGKQYVAEYV 581

```

RESULT 2
DHET-ACEAC
ID DHET-ACEAC STANDARD; PRT; 742 AA.
AC F18278;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RX MEDLINE=9255070; PubMed=2727242;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RA Yano K.,
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti.";
RL J. Bacteriol. 171:3115-3122(1989).
RN 3D-STRUCTURE MODELING.
RX MEDLINE=9528964; PubMed=7772016;
RA Cozier G.E., Giles I.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens.";

```

RL Biochem. J. 308:375-379(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.
CC -1- SUBUNIT: Tetramer of non identical chains (DEHYDROGENASE,
CC CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D90004; BA14058.1; -
DR PIR; J50326; J50326.
DR HSSP; 0924J7; 1FIC.
DR Interpro; IPR001479; Bac_POQ.
DR Interpro; IPR002372; Bac_POQ_repeat.
DR Pfam; PF01011; Bacterial_POQ.
DR PROSITE; PS00363; BACTERIAL_POQ_1.
DR PROSITE; PS00364; BACTERIAL_POQ_2.
DR PROSITE; PS00190; CYTOCHROME_C_1.
KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 742
FT ACT_SITE 343 343 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BINDING 649 649 HEME (POTENTIAL).
FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 742 AA; 81521 MW; 9C6C9268DAB825A CRC64.

```

Query Match 17.7%; Score 475; DB 1; Length 742;
Best Local Similarity 26.0%; Pred. No. 5.9e-28;
Matches 160; Conservative 87; Mismatches 202; Indels 166; Gaps 23;

```

OY 11 AGALALLAARPAQVPTVDELT--ANPAGEMISYGOENVRHSPLQITTNVQGLD 68
DB 23 AALPVAAYPARADGOGNGTEAIHADDHPEWLTSGRTYSEORSPIDQINRSNVGDLK 82
OY 69 LVNARGMOPKQVYV-TPILHGVYLANGVYQAIADAKTGLIWEHRRLP-NIA----- 122
DB 83 LIGYTLIDTRNGEGTPIVDGMYATTWMMKAVDADATGLIWDYDPRVGNIDKGC 142
OY 123 --TLN----- 126
DB 143 CDTVNRGAGYRWNGKVFMGTFDGRVLAADAKTKRVYAVNTIPADASLGKORSYVDGAVR 202
OY 127 VANGVYVASTCOYSPFGC--FVSGHDSATGEBELMRNFETIRAGEED----- 172
DB 203 VAKGLVYIGN--GAEERARGFVSAFPAETGKLWRVYTPVNNKNEPDHAAADNLIANKA 260
OY 173 -ETWNYDEAKWMT-----GAMQITYPVTNLVHGSTAVGPASETGRTGCTGLY--- 223
DB 261 YKTMGP--KGAMVROGGGGTVDLSLYDVPYSDLY--LAVG-----NGSPMNYKRSR 309
OY 224 --GTN-----TFEAVRPTGTGEIYWRHQTLPDRNWDCECTEFEMVTVNDVQPTSEBGLQSI 277
DB 310 GIGSNLEFLSIALKPEETBEYWHQETPMQDFTSVQOIMTIDLMPYK----- 358
OY 278 NPNMAGE-RVULTGVPCKTGTMMQDFAEGEIMAROTNYONMIESID-----EN 327
DB 359 -----GEMHVIHAP-KNGFEYIIDAKTGFIISGKNVYVNASGIDPRTGPIYNPALVTL 411
OY 328 GIYTVNEADILKEIDVEYDVCPTFLGGRDPSAALNDSGIYFIPLNVCY----- 378
DB 412 GLYTLNG-----KFWGI--PGLGAHNEMAMAYSPKTHLVYIPAHQIPFGYKNVGGF 463

```


QY 379 -----DMAVDEFTSMVYNTSKLPCKDMIGRIDALISTGRFLMSVERAANSP 434
 DB 464 KPHADSMVNGIDMTKNGLPDPE-ARTAYITADLGLWLANDPKMETWKIDKGPWNG 522
 QY 435 VJSTGGVYFNGGDRYPRALSOETGETLMOTRLATVASGQAISEYDGMQYVA----- 488
 DB 523 ILATGDLLEFOLANGEFHAYDATNGSDLYKFAQSGIAPMTYSVNGQYAVEVGNG 582
 QY 489 ----TAGGVSYSG 499
 DB 583 GYPISMGVYGRISG 597

RESULT 3
 DHET_ACEPO STANDARD: PRT: 738 AA.
 ID DHET_ACEPO
 AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DT Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 DE ADHA.
 GN Acetobacter polyoxogenes.
 OS Acetobacter; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Acetobacter.
 CC NCBL_TaxID=439;
 RX STRAIN-NB11028;
 RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RT Biochim. Biophys. Acta 1088:292-300(1991).
 RL -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POO AND HEME.
 CC -1- SUBUNIT: HETEROHEPIMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D00635; BAA00528.1;
 CC PIR: S14270; S14270.
 DR HSP: Q9Z4U7; 1f1G.
 DR IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR000345; CYC_heme_bind.
 DR Pfam: PF01011; Bacterial_POO_6.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR PROSITE: POO; Heme; Periplasmic; Membrane; Signal.
 DR OXidoreductase; 1; 35
 DR SIGNAL
 DR CHAIN 36 738
 DR BINDING 650 630
 DR BINDING 653 653
 DR METAL 654 654
 DR SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CR64;
 SQ

Query Match 17.6%; Score 474.5; DB 1; Length 738;
 Best Local Similarity 25.6%; Pred. No. 6,4e-28;
 Matches 151; Conservative 229; Indels 133; Gaps 15;

QY 10 SAGALLALPAFAQYTPVTDE-----LIANPPGEWISYSGONENYRHSPLQIT 61
 DB 17 TAGTICALISGATWASADDSOGATGEALIHADHPGNMFTYGRPTSDQKSPLDQINR 76
 QY 62 ENVGQOLYVARBMQKGYOV--TPLIHDSVMYLANPGVQAIQAIKATGDLIWEHRRLP 119
 DB 77 SNVGNLKIAYLIDLTNRQGEPTPLVDGVMYATINMSKMKVDAATGKLMSYDPKYPG 136
 QY 120 NIA-----TLN----- 125
 DB 137 NIADKGCDDIVNRGAAYNGKVFYFEDGRLIALDAKTKLWVSNTIPEAELEKORSY 196
 QY 126 -----IYANGVIYAGSNCOYSPGC--FVSGHDSATGEELRNFIPIPAGEED----- 172
 DB 197 IVDGAPRIAKGVIIIGN--GSEFGARGFVSARFADAEYKQVWRFVTPNPKNEPDAIDS 254
 QY 173 -----ETWGNDEYARMWTCG-----WGQIYDPVTMIVHGSYAVNGPASETOR 215
 DB 255 VLMNKAYQWTS-----PTGAWTRQGGGQGTWDSIYDDVADLVYLGVNGSGPMYKRYR 307
 QY 216 GTPGGTLYGTNRFAVRPDTEIWRHQTLPDRWDOCEIFEMAVTNVVDYQSTEMEGLQ 275
 DB 308 SEKGDNLFGLSIVALKRPTGEYVWHFOETPMDOQDFTSDQINTLDP1----- 357
 QY 276 SINPNATGERRVYLVGPKCTGMQFDEGEFLARDYVQNMIESID-ENGIVTNE 334
 DB 358 -----NGETRHVIYHARKNGEFTYIDAKTEFTSGNNYVVMASGLDKTRPIYNP 410
 QY 335 DALLKELDYVDYCPFLGGRDPSALNPDGSIYFIPLNVCYDMMAVDEFT-SMDVY 393
 DB 411 DALYTLTGKVEYIGPGLGSHNFAAMAFSPKGLVYIPAOVPELYTQVGFPIRHPDSW 470
 QY 394 NTS---NYTKLPFG-----KDMIGRIDALISTGRFLMSVERAANSPVJSTGGGV 442
 DB 471 NLGDMNKVGIIPDSPEAKQAFVDLKWIVAMPQQAEMWHDHKGPMNGGIIATGDDL 530
 QY 443 LFNGGTDYFRALSOETGETLMOTRLATVASGQAISEYDGMQYVAIAG 492
 DB 531 LFQGLANGEFHAYDATNGSDLYKFAADSGIAPPYTYLANGQYAVEVG 580
 RESULT 4
 DHET_GLUOX STANDARD: PRT: 757 AA.
 ID DHET_GLUOX
 AC 003542;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH
 DE subunit 1).
 GN ADHA.
 OS Gluonobacter oxydans (Gluonobacter suboxydans).
 OS Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluonobacter.
 CC NCBL_TaxID=442;
 RX STRAIN-IFO 12528;
 RX MEDLINE=97208225; PubMed=9055427;
 RA Kondo K., Horinouchi S.;
 RT "Characterization of the genes encoding the three-component membrane-
 RT bound alcohol dehydrogenase from Gluonobacter suboxydans and their
 RT expression in Acetobacter pasteurianus.";
 RT Appl. Environ. Microbiol. 63:1131-1138(1997).
 RL -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POO AND HEME.
 CC

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
 CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D86375; BAA19753.1; -
 DR HSSP; Q924J7; LF1G;
 DR InterPro; IPR001479; Bac_PQO.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 6.
 DR PROSITE; PS00363; Bacterial_PQO_1; 1.
 DR PROSITE; PS00364; Bacterial_PQO_2; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 DR Oxioreductase; PQO; Heme; Periplasmic; Membrane; Signal.
 KW SIGNAL 1 34
 FT CHAIN 1 34
 FT MOD_RES 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT ACT_SITE 342 342 PYRROLIDONE CARBOXYLIC ACID.
 FT BINDING 653 653 BASE (POTENTIAL).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82668 MW; 3989F90E3B947581 CRC64;

Query Match 17.2%; Score 461.5; DB 1; Length 757;
 Best Local Similarity 25.5%; Pred. No. 6.2e-27;
 Matches 156; Conservative 87; Mismatches 225; Indels 143; Gaps 21;

QY 7 LMSAGALALIAA--PAFAV--TPVDELLANPAGEMISYGOENYHRHSPLOTIT 61
 DB 16 LLSCAALALSAVPAFAFQEDGTATITSSDNGHP--DMLSYGRSYSEQYSLDDINT 74
 QY 62 ENYGOLVWARGMOPGVV--TPLLHDGVWYLANPVDVIAQIDAKGDLWEHRQLP 119
 DB 75 ENVGKLTLMHYDLDITNRQEGSTPLIVGVWATTNMSKALDAAGKLLMSYDPVPG 134
 QY 120 NTA-----TLNIVA--NGVIVAGS----- 136
 DB 135 NINADRCCTVRGAAYNGKVFETEDGRLIALDAKGLVMSVYTIPEAQLGHORST 194
 QY 137 TGOYSP-----FGC--FVSGHDSATGEELRNRYETIRAGEEGD----- 172
 DB 195 TYDGAAPRIAKKGVLLINGNGAEFGARFVSAPDAETSKLDRFTVNPENKPDGAASDDI 254
 QY 173 -----ETWGNDEYARMMWTG--AMGOITDPVTNLVHGSAVAVGASSTOGTGGTLY 223
 DB 255 LMSKAYPTMGKNGAMKQGGGGVYWSLVYDPTDLYLVGCVGSPWNNKFSBEKGNL 314
 QY 224 GTNRFAVRPDTEIYVHRQTLPRDNDQCEPEMVAIVNDVQSTEMEGLOSNPNAT 283
 DB 315 FLGSIIVAINDPDTRKYVHFQETPMDEMVTYVQQLITLMPV-----N 357
 QY 284 GE--RRVLTGVPCGTMMQDFATGTEFLMARDTNTQNNIESIDE--NGLVTVNEDAILKEL 341
 DB 358 GEMRHVIVAP--KNGFFYIIDAKTGKFTYIGKPYEENANGLDPVYGPANVPALMTLT 416
 QY 342 DVEYDCPFLGLGRDWPASALNPDGSIYFIPNNV-----CYDMAVDOEFT 388
 DB 417 GKPLDIPBELGHNFAAAYSPKTKLYITPAQDIPLLYDGQKGFAYADANNGLIDNN 476
 QY 389 SMDVYTSVWTLPKPKMIGIDAIDISTGRTL--WSVERAAAYS-----PVLST 438
 DB 477 KIGLPDNDPEHVAKKDF-----LKVLAGWIVAMPPEKMAFPIINHGPNGLLAT 530

QY 439 GGVLENGGTDTRYFRALSOETGETLMQRLTNVASGAISYVDGMOYAI----- 489
 DB 531 AGNVIFQGLANGFEHAYATATNNDLXSPQASAIAPPYTANGKQYVAEVGMCITP 590
 QY 490 -AGGGSYSYG 499
 DB 591 FLXGVARISG 601
 RESULT 5
 DHML_METEX STANDARD; PRT; 626 AA.
 AC P16027;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 GN alpha subunit) (MDH).
 OS MOXF.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID:408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AM1 / NCIB 9133;
 RX MEDLINE=90337342; PubMed=2116368;
 RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
 RT "Nucleotide sequence of the Methylobacterium extorquens AM1 mofx and
 RL gene 90:173-176(1990)."
 RN [2]
 RP SEQUENCE OF 28-53.
 RC STRAIN-AM1 / NCIB 9133;
 RX MEDLINE=89350892; PubMed=2504152;
 RA Nunn D.N., Day D., Anthony C.;
 RT "The second subunit of methanol dehydrogenase of Methylobacterium
 RL extorquens AM1.";
 RN Biochem. J. 260:857-862(1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=95384759; PubMed=7656012;
 RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
 RT "The active site of methanol dehydrogenase contains a disulphide
 RL bridge between adjacent cysteine residues.";
 RN Nat. Struct. Biol. 1:102-105(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
 RX MEDLINE=95253818; PubMed=7735834;
 RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
 RT "The refined structure of the quinoprotein methanol dehydrogenase
 RL from Methylobacterium extorquens at 1.94 A.";
 RN Structure 3:177-187(1995).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO.
 CC -1- SUBUNIT: THE HOMOZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31108; AAA25380.1; -
 DR PIR; S07908; S07908.

DR PIR: J00706; J00706.
 DR HSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 DR Oxidoreductase; P00; Signal; Methanol utilization; Periplasmic.
 KW SIGNAL 1 27
 FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330
 FT SEQUENCE 626 AA; 68434 MM; 649880A0FD2AD34C CRC64;
 SO
 Query Match 16.2%; Score 435.5; DB 1; Length 626;
 Best Local Similarity 24.6%; Pred. No. 4.3e-25;
 Matches 148; Conservative 91; Mismatches 208; Indels 155; Gaps 23;
 OY 10 SAGALALLA-AP-AFAOYTPVTDLIANPAGG-WISYGOENYRHSPLTQITTEVNGQ 66
 DB 7 SVSALMALALAPALSSGAYANDKVELSKSDNWMPGKNDSNNFSLKQIKGNVKK 66
 OY 67 LQIYW--ARGMOPGVQVYPLIHGVMY-----LANPGV----- 99
 DB 67 LRPAMFTSTGLNGH-EGAPLVYDGKMYHTSPNNTPALGLDDPGLIMODKPKONPAA 125
 OY 100 -----IOAIDAKTGDLIWEHRQLPNIATLN 125
 DB 126 RAVACCDLVNRGLAYWPGDGKTPALIKTQLDGNVALNAETGEYWK----- 173
 OY 126 IVANGVIYVASTCOYSPF-----GC-----FVSGHDSATGEELMRYNFI----- 164
 DB 174 -VENSIDIKGSTLTITAPVYVKDKVIIGSSGAEIGRYLTAIVYKGEQVRAVATGPDK 232
 OY 165 -----PRAGEE--GDETWNDEYARWMTGA--WQOITDPVNTLVHGSTAV 207
 DB 233 DLLLADDFNKNHAYGOKGLGTATWEGD--AMKIGGTWNGVAYADPGTNLITFGTNP 289
 OY 208 GPASSETGCTPGGTLTGNTREAVRPDGEIYWRHQLPRDNWDOCTEFEMVTVNDVOP 267
 DB 290 APNNTETMR--PGDNKM--TWITIGRADDTGEAKFGYOKTPHEMPO-----YAGVAVMM 338
 OY 268 STEMEGLQSIIPNAATGERVLTGVPCKTGTVMOPDATTGEFLMAR--DTNYONMIESI 324
 DB 339 LSEKAD-----KDGKARKLLTHPRDNGIYTLTDRDGLVSNKIKDDT--VNVFKSV 388
 OY 325 DENGIYVNEDELKELD--VEYDVCTPLFGGRDWPSSALNPDGSIYFPLNNVCYDMMAV 383
 DB 389 DLTGQAPVDEPGETRMDLAKDICPSAMGYHNOCHSDYDPKRELEFPGINHCIDMEPE 448
 OY 384 DOEFMSDVNTSVTKLPPEK-----DMIGRIDAIIDTSGRTLMSVERAANYSPVLS 437
 DB 449 MLPYRAGGFVAGATILNMPGPKGRONYEGLQIKAYNAITGKYKKEKERRAVVNGTMA 508
 OY 438 TGGVILFNGGDRYFRALSOETGETLWOTRLATVYASQALSYEVDSQYVALAAGGVSYG 497
 DB 509 TAGDILVFYGTGLDGLKANDSDTGLMKFKIPSGALIGYPMYTHKQYVAL-----YIG 563
 OY 498 SG 499
 DB 564 VG 565
 RESULT 6
 DHML_METOR STANDARD; PRG: 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large

DE alpha subunit) (MDH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OC NCB1-taxid=410;
 OX (1)
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
 RP STRAIN=ATCC 27866 / DSM 760 / NCIB 11278;
 RX MEDLINE=89008094; Pubmed=2459109;
 RA Machlin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 RT Methylobacterium organophilum XX methanol dehydrogenase structural
 RT gene".
 RT J. Bacteriol. 170:4739-4747(1988).
 RL J. CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: P00.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL P00 DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M26629; AAA50289.1; -
 DR HSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 DR Oxidoreductase; P00; Signal; Methanol utilization; Periplasmic.
 KW SIGNAL 1 28
 FT CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442 BY SIMILARITY.
 FT ACT_SITE 330 330 BASE (POTENTIAL).
 FT SEQUENCE 626 AA; 68677 MM; 8768F6B8371E5DFF CRC64;
 SO
 Query Match 16.2%; Score 435.5; DB 1; Length 626;
 Best Local Similarity 24.6%; Pred. No. 4.3e-25;
 Matches 148; Conservative 91; Mismatches 208; Indels 155; Gaps 23;
 OY 10 SAGALALLA-AP-AFAOYTPVTDLIANPAGG-WISYGOENYRHSPLTQITTEVNGQ 66
 DB 7 SVSALMALALAPALSSGAYANDKVELSKSDNWMPGKNDSNNFSLKQIKGNVKK 66
 OY 67 LQIYW--ARGMOPGVQVYPLIHGVMY-----LANPGV----- 99
 DB 67 LRPAMFTSTGLNGH-EGAPLVYDGKMYHTSPNNTPALGLDDPGLIMODKPKONPAA 125
 OY 100 -----IOAIDAKTGDLIWEHRQLPNIATLN 125
 DB 126 RAVACCDLVNRGLAYWPGDGKTPALIKTQLDGNVALNAETGEYWK----- 173
 OY 126 IVANGVIYVASTCOYSPF-----GC-----FVSGHDSATGEELMRYNFI----- 164
 DB 174 -VENSIDIKGSTLTITAPVYVKDKVIIGSSGAEIGRYLTAIVYKGEQVRAVATGPDK 232
 OY 165 -----PRAGEE--GDETWNDEYARWMTGA--WQOITDPVNTLVHGSTAV 207
 DB 233 DLLLADDFNKNHAYGOKGLGTATWEGD--AMKIGGTWNGVAYADPGTNLITFGTNP 289
 OY 208 GPASSETGCTPGGTLTGNTREAVRPDGEIYWRHQLPRDNWDOCTEFEMVTVNDVOP 267

Db 290 AFNMTMR--PQDNKM-TMTIFGRADTGEAKRGYQKTPHEDWDYAGVNM-----P 339
 QY 268 STEMEGLQSIINPNATGERRVLTGVPCKTGMQFPAEERGEFLMAR--DTYVQWMIISI 324
 Db 340 SEQNK-----KDGTRKLTHPRNGIYVTLTRTGALVSANKLDDT--VAVFKTV 388
 QY 325 DENGIVYNEDAILKEID-VEDVCPFLGGRDMPSAALNPDGSIETPLNNVCYDMMAV 383
 Db 389 DLKGTQVPRDEYGTMRDHLAKDVCPSAMGYNHGSDTDPKRELFEMGINHICDMEPE 448
 QY 384 DQETSMQVNTSVNLTLPCK-----DMGRIDALIDISTGRTLSVERAANYSPVLS 437
 Db 449 MLPYRAGQFVGAATLNNPCKKDRQNYEGLOIKAYNAITGSGYKMEREFVAVGGLA 508
 QY 438 TGGGVLENGTDRFALASQETGELTMTQRLATVAGSQAISYEDQMGYVLAAGGVSYG 497
 Db 509 TAGDLVFTGLDYLAKARDSDTGLMKFKIPSGAIGVPMYTHKGTQYVAI-----YVG 563
 QY 498 SG 499
 Db 564 VG 565

RESULT 7

XOXF_PARDE

STANDARD: PRT; 600 AA.

ID XOXF_PARDE

AC P29368;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-)

GN XOXF.

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Paracoccus.

OX NCBI_Taxid-266;

RN [1]

RP SEQUENCE FROM N.A.

RA Harms N.;

RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 380-600 FROM N.A.

RC STRAIN-PD 1235;

RX MEDLINE-92041583; PubMed-1657873;

RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,

RT Stouthamer A.H.;

RT "Isolation, sequencing, and mutagenesis of the gene encoding

RT cytochrome c551 of Paracoccus denitrificans and characterization of

RL the mutant strain".

RL J. Bacteriol. 173:6971-6979(1991).

CC -1- COPACITOR: P00 (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL P00 DEHYDROGENASE FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC

CC EMBL: U34346; AAC44555.1;

DR EMBL: M75583; AAA25574.1;

DR PIR: A41378; A41378.

DR HSSP: P38539; AAAH.

DR InterPro: IPR002372; Bac_P00_repeat.

DR Pfam: PF01011; Bacterial_P00; 7.

KW Oxidoreductase; P00; Signal.

FT SIGNAL 1 21

FT CHAIN 22 600

FT ACT_SITE 318 318

FT SEQUENCE 600 AA; 65159 MW; DCA996F1BC05A3CE CRC64;

SO

Query Match 15.98; Score 426.5; DB 1; Length 600;
 Best Local Similarity 25.88; Pred. No. 1.9e-24;
 Matches 156; Conservative 82; Mismatches 214; Indels 133; Gaps 26;

QY 10 SAGALALLAAPAFVQVPTDELALNP-----PAGEATISGQNDENRHSPLQITTE 62
 Db 6 NCACIALIMSGTAA-----LANEQRGRORQAPQAWIOMGVANTRSTSLQINKD 56
 QY 63 NVGQLOLVW--ARGMQRGVQVPLIHGVYLANP-GDVIGAIQDAK--TGDLMEHR-RQ 117
 Db 57 NVKDLRAVMTSTVLNGH--EGSPVIGVMVYHPPFRRVPAALDNGKILRWYEQQ 115
 QY 118 LPN-IATLNT-----VANGYIVAGSTCOYSPGCFPSGSDATGELKRNTP-IPRA 167
 Db 116 DPNVIAVMSQDYVYRGISYADGMIILGQA-----DTIVVALDNTSGEYKSTRIGDGI 169
 QY 168 GE-----EGDETW-----GNDYEA----- 181
 Db 170 GETLTATVPVKKVLYVSGEYVGRGMALNLTDSSEAKMAMSTGPPDELLVDPETT 229
 QY 182 -----RWMTGA-----WQITVDPVYVNLVHGVSTAVGASFGRTG 219
 Db 230 THLGKPIGADSSLSNMEGDQWQIGGCTIGWFSYDPLNLVYIGTGPSTWNSQR-PG 287
 QY 220 GTLYGTNTRFAVRPDTGETVYRHOITLPDNDQDECTEMAVTN--VDVQSTEMEGLQSI 277
 Db 288 DNKW-SMTIMARDADTGNAKFFQMTPHDENDVDGVNEMITLQVTDGQ----- 335
 QY 278 NNANATGERVLTGVPCKTGMQFPAEERGEFLMARDTYVQWMIISI-----SIDENGIYV 332
 Db 336 -----ERKLTTHFD--RNGLAYTMDRETGELLVAE--KYDPVYVNTTGVYDMDPNS-ET 384
 QY 333 NEDALINKELDYED-----YCPFLGGRDMPSAALNPDGSIETPLNNVCYDMMAVD 384
 Db 385 GRPVAEYVSTAQNGEDENTTGCIPALGTQDQPPAFSPKTNLFYVPPNHCMDYEPFR 444
 QY 385 QETSMQVY--NTSNVTKLPCKDMTIGRIDALIDISTGRTLSVERAANYSPVLS 442
 Db 445 VAYTAGQPIYGATLSMYPANSHSGMGNFIAWHNTTGELKNSVYRPSVMSGALATAGDV 504
 QY 443 LFNQGTDRFALASQETGELTMTQRLATVAGSQAISYEDQMGYVLAAGGVSYG 497
 Db 505 VFYGTLEGYLYKPYDAGTGBELKFKFTPSGIIIGNVATYEHGKQYVGLISGVGWAGIGLA 564
 QY 498 SGLN 502
 Db 565 AGLTN 569

RESULT 8

DHML_PARDE

STANDARD: PRT; 631 AA.

ID DHML_PARDE

AC P12293;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large

DE alpha subunit) (MDH).

GN MOXF.

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Paracoccus.

OX NCBI_Taxid-266;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.

RX MEDLINE-87307969; PubMed-3114231.

RA Harms N., de Vries G.E., Maurer K., Hoogenblik J., Stouthamer A.H.;

RT "Isolation and nucleotide sequence of the methanol dehydrogenase

RT structural gene from Paracoccus denitrificans.";

RL J. Bacteriol. 169:3969-3975(1987).

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +

```

CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M17339; AAA8366.1; -.
CC DR HSSP: P38539; AAAH.
CC DR InterPro: IPR001479; Bac_POQ_repeat.
CC DR InterPro: IPR002372; Bac_POQ_repeat.
CC DR Pfam: PF01011; Bacterial_POQ_7.
CC DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC FT SIGNAL 1 32 METHANOL DEHYDROGENASE SUBUNIT 1.
CC FT CHAIN 33 631 BY SIMILARITY.
CC FT DISULFID 135 136 BY SIMILARITY.
CC FT DISULFID 418 447 BY SIMILARITY.
CC FT ACT_SITE 335 335 BASE (POTENTIAL).
CC FT SEQUENCE 631 AA: 69799 MW: 0934DC93F5730B CRC64;
SQ
Query Match 15.8%; Score 424; DB 1; Length 631;
Best Local Similarity 24.9%; Pred. No. 3.1e-24;
Matches 153; Conservative 85; Mismatches 201; Indels 176; Gaps 26;
4 TSLMASAGATALLAAPAFQVTPYDEL-LANPPAGEMISGONENRHSPLTOIT 60
12 SSLMAVAVKAVLUTAPATA--NDQVLADPA-NMVMGRDINAQNEEMDIN 65
61 TENVGQQLVW--ARGM-----OQGV--QVTP- 84
66 KENVKQLRPAPSWFTGLHGEPTLVGDMFTPTPNTTFALDLPFGKILMOKPK 125
85 -----LHIDGVMTLANGD-----VIOALDAKGDLMHEHRQL 118
126 ONPTARTVACDGVNRLAYW--FGDOVAPLIFRTQDGHVAMDAETGERM----- 177
119 PNIAITLVANGVYAGSTCOYSPF-----GC-----FVSGHDSATGEIWRN 161
178 -----IMENSDIKVSTLTIAVYKIDLVGSGCAELGVAGVYAVDVSGEMRRA 230
162 YFI-----PRAGE--GDETWGNDYBARMTGA--WQGITDPYTNLV 200
231 FAIGPEDELLAEDFNAPNPHGOKNLETWEGD--AMKIGGTWGWYAVDPEVDLF 287
201 HYGSTAVGASSTGPGTLYGNTREFAVAPDTELIVRHQTLPPDNNDQCTEEMV 260
288 YVSGNPPAPNEMR--PDKNW--TMAIMGREATGEEKFAVYQTPDEMD-----Y 336
261 TNDVOPSTMEGLSINPAATGERVLTGVCPTGTMTGMPDAETGEFLWARDNTYOM 320
337 AGVAVVWMLSEQEDKO-----GOMRKLTPDRNGVITLDRNGDLISA----- 380
321 IESIDE--NGIVYNEDALIKELDEY-----DVCPTLGGMDPMSALNDGSIY 369
381 -DKMDVYVWYKEVQDGTGPRDPEFGFMHAKRDICPSAMGYHNOCHSDYPERKVF 439
370 FIPLNNVCYDMAVDQFTSMYVNTNSVTKLPGRDM-----IGRDAIDISTGRTLS 424
440 MGINHICMDWEPMLPFRAGQFVGATLTYGPKATLBRAGAGIKAYDAISEMKWE 499
425 VERAANVSPLSTGGGVLENGTDRYFRALSOETGETLMQTLAIVASGOAISEVDGM 484

```

```

Db 500 KMERFVWGCTAATAGLTFYVTLDFIKARDSDTGLMKFKLPSCVYGHPTKYNHGR 559
QY 485 QYVIAAGGVSYSGS 499
Db 560 QYVAL-----MYGVG 569
RESULT 9
EXAA_PSEAE STANDARD; PRT: 623 AA.
ID EXAA_PSEAE
AC 092437;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEHD).
GN EXAA OR P1982.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Braden L.L., Coulter S.N., Folger K.R., Kas A., Landig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen.";
RL Nature 406:959-964(2000).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=99173751; PubMed=10075429;
RA Schobert M., Goerlich H.;
RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
RT dehydrogenase.";
RL Microbiology 145:471-481(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
RC STRAIN=ATCC 17933;
RX MEDLINE=20202376; PubMed=10736230;
RA Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
RA Goerlich H.;
RA "X-ray structure of the quinoprotein ethanol dehydrogenase from
RA Pseudomonas aeruginosa: basis of substrate specificity.";
RL J. Mol. Biol. 297:961-974(2000).
RN [5]
RP FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
RN ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
RN -1- COFACTOR: POQ AND CALCIUM.
RN -1- SUBUNIT: HOMODIMER.
RN -1- SUBCELLULAR LOCATION: Periplasmic.
RN -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ009858; CAA08866.1;
 DR EMBL: AE004624; AAC05370.1;
 DR EMBL: AF068264; AAC79657.1;
 DR PDB: 1FLG; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 DR Oxidoreductase: PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 DR Complete proteome: PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 FT SIGNAL
 FT CHAIN 1 34
 FT DISUFID 139 140
 FT SQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;
 QUNOPROTEIN ETHANOL DEHYDROGENASE.

Query Match 15.4%; Score 413; DB 1; Length 623;
 Best Local Similarity 23.3%; Pred. No. 2.1e-23;
 Matches 140; Conservative 93; Mismatches 245; Indels 122; Gaps 20;

OY 3 PSLIMASGALAL--LAAPAFQVPTDELALN--PPAGEMISGQOENYRHSPLT 57
 DB 9 PGLLRPLSLHCAFAVALGSAALANDVEMDIANDKTTGYLDYGMGTNAORWSPLK 68
 OY 58 QTTTENVOLOLVARGM---OPKQVQVPLHGVMTLANGDVIOAIDAKTGLIME 113
 DB 69 QVNAADNVKFLTPAMSYSGDEKQGO--ESQAIYSDGIYVYASRLFALDAKTKRLMT 127
 OY 114 HRRLP.N-----IATL--NIVA-----NGVY-----AGST 137
 DB 128 YNHRLPDIPRCDDVNRGAIIYGDKVFETLDASVALNKTGKVKKKFADHAGCT 187
 OY 138 QYSP-----FCGF--VSGHDSATGEELMRNFIPR-----166
 DB 188 MTGAPITVKGKTKGKVLHSGSGDEFGVGRFARDPTDTEIMRFEVGHNGRLNGK 247
 OY 167 -----AGEGEETWANDYEA-----RMTGA-----NGQITTPVNLVHYSGTANGPASET 213
 DB 248 DSTVTGVAAPWDDRNPSPTKVESWSHGSGAPWQSAFLAETNITITVAGNPGPMTW 307
 OY 214 QRGTPGG-----TLGVTFRFVRPDGELVNRHQLPRDNMDQCTFEMATNVADVP 267
 DB 308 APTAKGNPHDYDSLX--TSGOVGVDPSGGEYKWFYQHTPNDAMDFSGNNELVLFDKRAD 366
 OY 268 STEMEGLASINEN-----AATGERRVLTGVPCKTGTW--QFDAETGEFLMADTYQNM 320
 DB 367 GRIIVATAHADNGFEVYVDRSNGKLNAPFPVDNITWASHIDLKGT-----RVERREGQ 421
 OY 321 IESIDENGIIVNEDALIKELDEYDVCPTFLGRDMPFSAALNPDGSIYFIPLNVCYDM 380
 DB 422 RPLPPEP-----QKHKAIVEYSPFLDGKNNMNPAYSDPTGLYYVPAHMKEDY 471
 OY 381 MAVQOEFSMDVYNTSVTLRPGKDMIGRIDALIDISTGRITMSVERAANSPLYLSGG 440
 DB 472 WTEEVSTTKSAIYGMGFRITKRYDDHVSILRAMDVSQKVVMEKEHPLMAGVLAAG 531
 OY 441 CVLFNGSTDRYFRALSOGETLQTRLATVASGAISTEVDGMYVAIAGGVSYSG 500
 DB 532 NLVFTGTGDYFKAFDAKSGKELMKFQJGSGIVSPITWEDGEOYLGYT---VGYGAV 588

DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha
 subunit) (MEDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 ON NCBI_Taxid-17;
 RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA MEDLINE=9405969; PubMed=8241148.
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RT Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RL methanol dehydrogenase."
 RL Biochemistry 32:12953-12958(1993).
 RP [12]
 RX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RA MEDLINE=93054513; PubMed=1331050.
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution."
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1 CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.
 CC -1 Cofactor: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1 SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1 SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 DR PDB: 4AAH; 08-DEC-96.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
 KW Calcium.
 FT DISUFID 103 104
 FT ACT_SITE 379 408
 FT ACT_SITE 297 297
 FT HELIX 2 9
 FT TURN 11 12
 FT TURN 14 14
 FT TURN 17 18
 FT TURN 21 22
 FT STRAND 26 27
 FT TURN 34 36
 FT TURN 37 39
 FT STRAND 41 47
 FT STRAND 49 61
 FT TURN 62 63
 FT STRAND 64 68
 FT TURN 71 73
 FT STRAND 75 79
 FT TURN 80 81
 FT STRAND 83 84
 FT TURN 85 90
 FT HELIX 96 101
 FT TURN 103 104
 FT TURN 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 136 142
 FT TURN 146 148
 FT TURN 149 149
 FT STRAND 151 151
 FT TURN 157 159
 FT TURN 160 161

BASE (POTENTIAL).

wallick-934-125.pcp.isp

Fri May 24 11:27:26 2002

FT STRAND 162 166
 FT STRAND 168 168
 FT STRAND 170 172
 FT STRAND 173 173
 FT STRAND 177 182
 FT STRAND 183 185
 FT STRAND 188 193
 FT STRAND 198 202
 FT STRAND 203 203
 FT STRAND 205 210
 FT STRAND 212 213
 FT STRAND 219 222
 FT STRAND 223 223
 FT STRAND 226 231
 FT STRAND 241 243
 FT STRAND 244 247
 FT STRAND 248 252
 FT STRAND 261 263
 FT STRAND 269 270
 FT STRAND 273 278
 FT STRAND 279 281
 FT STRAND 284 289
 FT STRAND 305 311
 FT STRAND 312 313
 FT STRAND 314 322
 FT STRAND 324 325
 FT STRAND 327 332
 FT STRAND 333 333
 FT STRAND 338 343
 FT STRAND 346 347
 FT STRAND 351 355
 FT STRAND 356 359
 FT STRAND 365 367
 FT STRAND 370 370
 FT STRAND 371 372
 FT STRAND 373 373
 FT STRAND 375 378
 FT STRAND 382 383
 FT STRAND 390 390
 FT STRAND 392 394
 FT STRAND 395 398
 FT STRAND 399 405
 FT STRAND 407 413
 FT STRAND 420 421
 FT STRAND 427 433
 FT STRAND 435 436
 FT STRAND 438 439
 FT STRAND 444 450
 FT STRAND 451 454
 FT STRAND 455 462
 FT STRAND 469 469
 FT STRAND 471 473
 FT STRAND 474 476
 FT STRAND 477 481
 FT STRAND 483 484
 FT STRAND 486 491
 FT STRAND 492 495
 FT STRAND 496 502
 FT STRAND 512 516
 FT STRAND 517 518
 FT STRAND 519 526
 FT STRAND 530 533
 FT STRAND 534 538
 FT STRAND 539 539
 FT STRAND 543 545
 FT STRAND 546 552
 FT STRAND 553 555
 FT STRAND 556 558
 FT STRAND 559 559
 FT STRAND 565 570
 FT STRAND 571 AA: 62449 MW: OBE94EA5A2DAB1E1 CRC64:

Query Match 14.4%; Score 388.5; DB 1; Length 571;
 Best Local Similarity 23.9%; Pred. No. 1.3e-21;
 Matches 136; Conservative 78; Mismatches 199; Indels 155; Gaps 22;

QY 38 AGEWISYGOENRHSPLTOITTEWGOLOLV--ARGMPKQVOTPLIHGVMY--- 92
 DB 11 AGAMPITAGGYSOHNSPLAQINKSVKVKAKMSFSTGVINGH-EGAPLVIGDMKTVHS 69
 QY 93 -----LANPGDVI----- 129
 DB 70 AEPNNITVLANLNDPGKIVMOHKPRKODASTKAVWCDDYVDRGLAGAGIVKQKANGHLLA 145
 QY 103 IDAKTGDLIMEHRROLNINATINIVANGYIVAGSTCQISPF-----GC----- 176
 DB 130 LIDAKTGKINMEVEYCDPV-----GSTLLOAFVAKDTVLMGSCGAEIVRG 187
 QY 146 FVSGHDSATGEELMR-----NYFIPRAGE--EGDETWGNDEYKAWMTGA 233
 DB 177 AVNAFELKTGELKWRFAFGTSDSVRLAKDPSANPHYGOGFGTGTWGD--AMKIGG 239
 QY 188 ---WGOITDPVTLNVHGSTAVGPASETQRTG-----TLGTNTRFAYAPDTGIV 285
 DB 234 GTMNGWYADPKLNLFYSGSNPAPMNETMR--PDKNKMTITWGRDL-----DTGMAR 297
 QY 240 WRHQTLPDRNMDOCTFEPMATNTVDOP--STEMEGLQS-INPNAATGERVVLGVCKTG 326
 DB 286 WGYOCTPDEWDEAGVNOVLID--QPVNAKMTPLSHIDRN-----G 326
 QY 298 TMMQFDEATGEFIWARDTN-YQNMIESIDENGIVYNEDALIKELDVE-XDVCPTFLGR 355
 DB 327 ILYTLNRENGNLIYAEKVADPAVNVFKVDKTGTVPYRDEPFATRDHGTNICSANGFH 386
 QY 356 DWPSAALNPDGSIYFIPLNANCYDMAVDOETSMQVYTSVNTLP---PGKDMIGRI 411
 DB 387 NOGVDSYDPESRTLYAGLNHCMDPEFMLPYRAGOFEGATLAMPFGNPGTKKEMQOI 446
 QY 412 DAIDISTGRTLMSVERAANSPVLSGGGVLFENGCTDRYFRALSOETGETLMOFRLATV 471
 DB 447 RAPDLTGTAKAKWTKEKFAWAGTLYTGGLVWATLDGLKALDKRDEKELNFKMPSG 506
 QY 472 ASGQATSYEVDGMOYVAIAGGVSYSG 499
 DB 507 GIGSPMTYSPFKQYI-----GSMWGVG 529

RESULT 11
 ID DHG_ECOLI STANDARD; PRT: 796 AA.
 AC P15877:
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update) (EC 1.1.99.17).
 DE Glucose dehydrogenase [pyrroloquinoline-quinone]
 GN GCD OR B0124.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12:
 RX MEDLINE-91035240; PubMed-2228962;
 RA Cleton-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
 RT "Cloning, mapping, and sequencing of the gene encoding Escherichia
 coli quinioprotein glucose dehydrogenase.";
 RL J. Bacteriol. 172:6308-6315(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RX MEDLINE-93123180; PubMed-8419307;
 RA Yamada M., Asaoka S., Sailer M.H. Jr., Yamada Y.;
 RT "Characterization of the gcd gene from Escherichia coli K-12 W3110

RT and regulation of its expression.";
 RN J. Bacteriol. 175:568-571(1993).
 RC SEQUENCE FROM N.A.
 RC STRAIN-K12 / M3110.
 RX MEDLINE-94261430; PubMed-8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Gregor J., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLGOL.
 RX MEDLINE-93286127; PubMed-8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of ubiquinone glucose dehydrogenase of Escherichia
 RT coli modelled on that of methanol dehydrogenase from Methylobacterium
 RL Blochm. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: X51323; CAA35706.1;
 DR EMBL: D12651; BAA02174.1;
 DR EMBL: D25652; CAA20298.1;
 DR EMBL: AB000122; AAC73235.1;
 DR PIR: JVO107; JVO107.
 DR HSSP: P98359; 4AAH.
 DR Ecogene: E810369; gcd.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1;
 DR PROSITE: PS00364; BACTERIAL_POO_2;
 DR Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT TRANSMEM 1 10 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 11 37 PROBABLE.
 FT TRANSMEM 38 40 PERIPLASMIC (PROBABLE).
 FT DOMAIN 41 58 PROBABLE.
 FT TRANSMEM 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).

FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 142 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TLSADAMP -> HIKRCHP (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SO SEQUENCE 796 AA; 86747 MW; D9EDC705A1289495 CRC64;

Query Match 12.6% Score 339.5; DB 1; Length 796;
 Best Local Similarity 22.5% Pred. No. 9, 2e-18;
 Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;
 QY 6 LMAS-----AGALLAAPAFQVTVPTDELLANPPAGISWISGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQRIKNTLSADATPAEA-ISPVDQ-----DMPAYGRNGGQRFSPKQ 185
 QY 59 ITTENVGQLOLWMA-----RGMGRKY--QVTPLEHGVVYLANPGDYIOAIDAKTGD 109
 DB 186 INADNVNHLKAWVFTGDKQPNDEGTTEVPIKVGDTLYLCTAHGRFLADASR 245
 QY 110 LIMEHRQL-----
 DB 246 EKHMHDELTNESFQHVTCRGVSYEAKETASPEVMADCPRIILPVNDRLAINAE 118
 QY 119 -----PNIATLNIYAN-----GVVAGS-TCQYS--PFGCFV 147
 DB 306 NGKLCETFAKGVNLMQSNMPTDRGLYEPTSPILTDKTIWAGSVTDFNSTRETSVI 365
 QY 148 SGHDSATGEELRNENFIRA-----GEGDETDGNDYEAEMWANGQOITPYPTNLY 200
 DB 366 REFVDVTGELLMA--PDGADPNALIPSDHEPTENS-----PNSAPRAITDAKLDY 416
 QY 201 HYGSTAVGPASETGRPGTGLVGTNTFRVAPDGEIYVHQTLPKPNMDQECTFEMV 260
 DB 417 YLPMGVTTPTDVGWGNRTPEQERYASSI-LALNATGKLAMSTQVHHDLMDMD- 468
 QY 261 TNDVDPSTMEGLSINPMAATGERRVLGVPCCTGTMMQFDAETGEPL----- 310
 DB 469 --LPAQPT-----LADITYN--GQVPIYAPATAGNIFVLDNRNELVPAPEKPVQ 518
 QY 311 WARDTNYQNMIESIDNGI-----VYVNDALIKEL--DYED- 346
 DB 519 GAAKGDIYVTPQFSLSFRPTKDLGADMGATMFDQLCRVHFQMRIGEIFTPPSBQ 578
 QY 347 ---VCTFLGGRDMPSSALNP-----D 365
 DB 579 GTLVPPGNLGMFEMGQISVDNREVALANMALPEVSKLIPKPGNPMRQPKAKGTGE 638
 QY 366 SGI-----YFIPLNVCYDMAVNDQETSMQVNTSVNTKLPKGMIRIDAIDIST 418
 DB 639 SGIQPGVGPYGVTLN-----PFLS-----PSELCKQAPAMYSIALDLKT 679
 QY 419 GRTLMVSVERAANYS-----PV-----LSTGCGVLFNGCT--DRFRALSDQ 458
 DB 680 NEVWAKRIGTIPDSMFPPVPVPEFNMGMPLGPISTAGNVLFIAATADNYLAVNMS 739
 QY 459 TGETLMTRLATYASGQA--ISYEVDGMQYVAT-AGGVSYS 498
 DB 740 NGKLMGRILP--AGQATPMTYEYVNGKQYVVISAGHGSGFT 780
 RESULT 12
 DHGA ACTICA STANDARD; PRT; 801 AA.
 AC P05463;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A (pyritolquinone-quinone) precursor

DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 AC Acinetobacter.
 NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LMD 79.41;
 RX MEDLINE=8828368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus.
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X0735; CA30222.1; -
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR OXIDOREDUCTASE: POQ; Transmembrane; Periplasmic; Signal.
 KW CHAIN 33
 FT SIGNAL 1 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 FT SEQUENCE 801 AA; 86956 MW; 2FAEA322EB5450D8 CRC64;
 SQ
 Query Match 11.7%; Score 315; DB 1; Length 801;
 Best Local Similarity 21.2%; Pred. No. 6.4e-16;
 Matches 148; Conservative 76; Mismatches 192; Indels 282; Gaps 30;
 QY 20 PAFQAVTVELLAPNAGEWISYQONENRHSPLTQITTEVNCQLOLVANRQCK 79
 DB 152 PETAADVGVAF-----SDMPAYGRTQAGVRSPLKQINDNVDLVAMT--LRTGD 202
 QY 80 V-----QVPLIHGVNLIANGVIOAIDAKTQDLIWEHNRRL----- 118
 DB 203 LKINDSGETTVQVPIKIGNNMFCTAHOQLAIDPATCKEWRFPKLTCKSFQHLT 262
 QY 119 ----- 118
 DB 263 CRGVWYDANNTTEFATSLQSKSSSTQCKRVFVNVNDGLVAVNADTKACIDFGNG 322
 QY 119 -----PNATLNIYANGVIAGS--TCOYS--PFGCVSGHDSAT 154
 DB 323 QVNIQEMPAYPAGYNPISPGIYV-----GSHVVIAGSVTINSKPESG--VIRGYDVT 377
 QY 155 GEELMRNYFIPRAGE-----EGDETGWNDYLRKMTGAMQGIYDPTNIVHYGSTAVG 208

DB 378 GKILW--VEDTGAADPNNAHGEETTFVHNS-----PNAAPLAYAKLIDV-YVPTGV- 427
 QY 209 PASFQRCGPGGTLYG-----TNTFAVPRDCEGLVHMHOTLPNDNDOCEFEH 258
 DB 428 -----GNP--DIWGGRTELKERYANSMALINASTGKLWNFQTHHLMQMDVPSOP 478
 QY 259 MVTNDVOPSTMEGLQINP--NNAATGERRVLTGV-----CKTG--TMMQF 302
 DB 479 SLADINKKAGQTPATVYLTGKGNAFVLDNR--NGQPIVPTKRPQVTKRGQTKGEF 536
 QY 303 DAEFGEF-----LMA-----RDNYNMIESIDENGIVTNE 334
 DB 537 YSKTQPSDLNLAPODKLIDKMDMGATMLDQMCVSKRLNIGITPPSENGTL----- 592
 QY 335 DALIKELDEYDCPTFLGSRDPMASALNPS-----GITYF----- 371
 DB 593 -----VFGNLGVFEWGMGMSVNPDRQVAVMNPDIGLFPVSRLLIPDPNRAQTK 640
 QY 372 -----PLNNVCYDMAVDOEFTSMQVNTSVTKLPKGDIMGRIDAIDISTGRTL 422
 DB 641 GASTEGGVQPMYGVVY-----GVEISAFSLPLGLPKQPMAGVYAGVDLTKTHEV 690
 QY 423 W-----SVERAANYSPVLSTGGVLFNGST--DRYFRALSOEFTGL 463
 DB 691 WKRRIGTIRDSLPLNLPOLPAVKIGVPGIGSISTAGVNMFGATQDNYLRANVTNGKKL 750
 QY 464 MQTRLATVAGQA--ISTEVDGMQYVAI--AGGGVSYGS 498
 DB 751 WEARLP--AGGQATPMYEINGKQVYVIMAGHGSFOT 786
 RESULT 13
 DHG_GLUOX STANDARD; PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconobacter.
 CC NCB1_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 of quinoprotein glucose dehydrogenase in Gluconobacter oxydans";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RU Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; OPKEX.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR00372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT_SITE 470 470
 FT VARIANT 788 788
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match
 Best Local Similarity 10.7%; Score 287.5; DB 1; Length 808;
 Matches 140; Conservative 82; Mismatches 229; Indels 251; Gaps 25;

13 ALALLAFAFAQVPTVDELTA-----NPPAGEWISTGQONENYRHSPLTQIT 61
 128 AVLAFAFASIFTDPHDISGLPTQIANSPADPDVPSSEHAYGRQAGDRMSPLNQINA 187
 62 ENVGQLOLV-----ARGMQEKV--QVPLIHGVYLANPQVQIAIDAKTGDLIW 112
 188 TNSVSLAKAAWHITKDKMANSNDPEQTEATPIEFNNLTLMGSLHOKLFAVAGATGNVWK 247
 113 EHRQL--PNIATLNI-----VANGVIV-----133
 248 VYDKLQINPFOHLCRGVSEFEPANAMDSGNAPLPDCAKSDILFVNDGRLEVDAD 307
 134 AGSTC-----QSPGFCV-----147
 308 TGTGSGFGNNGEIDLRVNPQPTTGGVYPTIDKLILANSATIDGSKAKSGA 367
 148 -SGHSDTSGEELM---RNYFTPRAGEEDFTWGDYDARMTGAMQIITYDPTNLVHY 202
 368 TQAFDYVTKRWVFDASNDPQQLPDESHPEHNSNSIVS---SYDANLNLVYI 422
 203 GSTAVPASETORGP---GGLYGTNRF---AVRPDTGELVNRHQTLPDNDMDQC 254
 423 PMGV-----GTPDQMGDRKDKSEREPAGIVALNADTGLAMFYQTVHHDLMDL 473
 255 TFEAMVTVNDVQPTSEMEGLQSLNPMAATGERRVLTGVPCKTGMQOFAETG-EFLMAR 313
 474 PQQPLVDTOKDGLVPAIYA-----PKTGDIVLDRTGKEIVPAP 517
 314 DT-----NYQNMIESIDENGIVTVNEDALIKELDV-----EAD 346
 518 ETPVPOGAPGDHSPQPMQ--LTLRPKNPLNSDITWGTFIDQMCSTYFTLTYE 574
 347 -----VCPITLGRDMPASALNDSGIYIPINNVCYDMAVQ-----385
 575 GPTTPPSLKSILFPGDLGFMWGLAVPQOVAFANISLPEVSQVLPSPGNPLMPE 634
 386 -----EFTSDVYNTSVTKLPQKDMI-----GRIDAIDISTGRTL 422
 635 ENAKGTGGEGTGLQHNIGIPYAVNLHFDLPVLLPFGIMPKRTPWGVAGIDTKTKVY 694
 423 WSYERAANVS-----PV-----LSTGGVLF--NGGIDRFPRALSGETG 460
 695 WQHRNCTLDSMKGSLPIPLPIKIGVPSLGGPLSTAGNLGFLTAMMDYIIRAYNLTTG 754
 461 ETLMTQRLTAVASGAISYEVDGMQYVAIAGGVSYSGSLNS 502
 755 KYLMODRLPAGAGQATPIYAIANGKXI-----VTAGGHS 790

RESULT 14
 ID OUIA.ACICA STANDARD; PRT; 809 AA.
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [Pyroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase);
 GN OUIA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsemore D.A., Ornston L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsemore D.A., Ornston L.N.;
 RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
 RT J. Bacteriol. 177:5971-5978(1995).
 RL [3]
 RP FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: Quinate + pyroloquinoline-quinone = 5-
 CC dehydroquininate + reduced pyroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: PQQ.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY: FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PROTOCATECHUATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sib.ch).
 CC
 DR EMBL: L05770; AAC37161.1; -
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR00372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE NEG.
 KW Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 FT TRANSMEM 147 147
 SQ SEQUENCE 809 AA; 88196 MW; 71P67CEBFA62BFCB CRC64;

Query Match
 Best Local Similarity 9.3%; Score 250.5; DB 1; Length 809;
 Matches 144; Conservative 70; Mismatches 242; Indels 223; Gaps 24;

walick-934-125.pwp.rsp

Page 14

Search completed: May 24, 2002, 10:30:52
Job time: 630 sec

Fri May 24 11:27:28 2002

walick-934-125.psp.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:15 ; Search time 93.8 Seconds
(without alignments)
925,836 Million cell updates/sec

Title: WALICK-934-125.PEP
Perfect score: 2689
Sequence: 1 MKPTSLMASAGLALALAP.....GMQVAVAGGVSYSGSLNS 502

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORNELLAE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1490	55.4	608	2	093RE9
2	520.5	19.4	698	2	09KH03
3	487	18.1	742	2	053362
4	482	17.9	601	16	09EWM8
5	481	17.9	601	16	09EWM8
6	479.5	17.8	708	2	046444
7	477.5	17.5	695	2	09F902
8	470.5	17.2	629	2	09AQ48
9	463.5	17.2	601	2	09AE95
10	461.5	17.2	601	2	09AE95
11	431.5	16.0	633	2	024759
12	428	15.9	623	2	09AGW3
13	419	15.6	695	2	091935
14	404.5	15.0	695	2	093460
15	391.5	14.6	573	2	059340
16	297.5	11.1	790	2	09X255

17	285	10.6	785	16	Q98KE6	Q98ALV	pancoea cit
18	280.5	10.4	786	2	P95466	P95466	pseudomonas
19	267	9.9	803	16	Q91115	Q91115	pseudomonas
20	258	9.6	777	16	Q92R83	Q92R83	rhizobium m
21	230	8.6	639	2	P77931	P77931	pseudomonas
22	225.5	8.4	644	2	O52551	O52551	pseudomonas
23	208.5	7.8	182	2	O32699	O32699	hyphomicrob
24	207.5	7.7	470	2	O30326	O30326	acetobacter
25	203.5	7.6	179	2	O32697	O32697	hyphomicrob
26	200.5	7.5	180	2	O32700	O32700	hyphomicrob
27	200.5	7.5	180	2	O32700	O32700	hyphomicrob
28	200.5	7.5	181	2	O32615	O32615	hyphomicrob
29	199.5	7.4	181	2	O32692	O32692	hyphomicrob
30	197.5	7.3	181	2	O32703	O32703	hyphomicrob
31	196.5	7.3	182	2	O32621	O32621	hyphomicrob
32	195.5	7.3	182	2	O32706	O32706	hyphomicrob
33	193.5	7.2	180	2	O32701	O32701	hyphomicrob
34	191.5	7.1	182	2	O32612	O32612	hyphomicrob
35	191.5	7.1	184	2	O32693	O32693	hyphomicrob
36	190.5	7.1	172	2	O33884	O33884	methylosinu
37	186.5	6.9	179	2	O32707	O32707	hyphomicrob
38	186.5	6.9	180	2	O32694	O32694	hyphomicrob
39	185.5	6.9	180	2	O32704	O32704	hyphomicrob
40	185.5	6.9	171	2	O33884	O33884	methylosinu
41	185	6.9	171	2	O32702	O32702	hyphomicrob
42	183.5	6.8	179	2	O32695	O32695	hyphomicrob
43	182.5	6.8	184	2	O918K8	O918K8	uncultured
44	181.5	6.7	185	2	O93K58	O93K58	methanotrop
45	176.5	6.6	172	2	O93K58	O93K58	methanotrop

ALIGNMENTS

RESULT	ID	Q93RE9	PRELIMINARY	PRT	608 AA.
AC	Q93RE9	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	ALCOHOL DEHYDROGENASE.				
GN	ADH.				
OS	Pseudoglucobacter saccharoketogenes.				
OC	Bacteria: Pseudoglucobacter.				
OX	NCBI_TaxID=133921;				
RN	SEQUENCE FROM N.A.				
RP	SRRAIN-IFO 14464;				
RC	Sribbata T., Saito Y.;				
RA	"Alcohol dehydrogenase."				
RT	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
RL	EMBL: AB046580; BAB62258.1;				
DR	EMBL: AB046580; BAB62258.1;				
SQ	SEQUENCE 608 AA; 65101 MW; OACEC97AE1BA570 CRC64;				

Query Match	Score	Length	DB 2;	Length	608;
Best Local Similarity	55.48;				
Matches	291;	Conservative	63;	Mismatches	138;
Indels	90;	Gaps	6;		
QY	10	SGALALIAAPFAO-----VTPYDELLANPAGEWISYSGONEN 50			
DB	16	STALIASGPFAPFHDNMAAEPKAGQSAIENFOPVADLAKNPAMWPLIRNGYOG 75			
QY	51	YRHSPLTQITTEVNGQDLVWARGMPGVOTPLIDGVYLANPGVDIQAIDATGDL 110			
DB	76	WGVSPLDQINKNVGDLQVMSRTMEPSNGGALAVNGVIFLGTNDVIOAIDGKGL 135			
QY	111	IMEHRQOLPN-----IATLN----- 125			
DB	136	IMEHRQOLPN-----IATLN----- 125			
QY	126	YANGVIYAGSTQYSPGCVSGHDSATGEELMRYNFIIPRAEBDEET 174			

```

Db 196 VEEGVANSSGPIVNDGVVIASTQCFSGCYVGTDAESGEELMRTFIPRGEEDT 255
OY 175 WGN-DYEAEMWTGANGQITTPYVNLVHYSTAGPASEFORPTGCTLYGNTREAVP 233
Db 256 WCGAFYERMMWTGAMGOITTPYDDELIVYSTAGPASEYOROTEGTLAGNTREAVP 315
OY 234 DTGEIWMHQTLPBNDNDQECTEMWYTNVDPSTEMEGLOSINPMATGE-RVLTGV 292
Db 316 KTGEVYMKHOTLPBNDNDQECTEMWYTNVDPSTEMEGLOSINPMATGE-RVLTGV 292
OY 293 PCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
Db 376 PCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
OY 353 GGRDMPGSAALNPDGSIYFPLNNVCYDMMAVDOEFTSMOYNTSNVTKLPKGMIGRID 412
Db 436 GGRDMPGSAALNPDGSIYFPLNNVCYDMMAVDOEFTSMOYNTSNVTKLPKGMIGRID 412
OY 413 AIDISTGRITLWSEVERAANYSPVLTGGGVLENGGTDRYFALSOETGELTNGTRLATYA 472
Db 496 AIDALGETKMSYETRALYDPEVLTGGDLVFGVIGIDRFDALDAESGKEVWSTRLPGAV 555
OY 473 SGQALSYEDGMOYVAIAGG-----VSYGSLN 501
Db 556 SGTTSYSIDGRQYVAVVSGSLGPTFGPTTDPVDSASGAN 597

RESULT 2
O9KH03 PRELIMINARY; PRT; 698 AA.
AC O9KH03;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE TERAHYDROFURORYL ALCOHOL DEHYDROGENASE
OS Alkaligenes eutrophus (Ralstonia eutropha)
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OC NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21122557; PubMed-11222593;
RA Zarit G., Schrader T., Andreesen J.R.;
RT Tetrahydrofurfuryl Alcohol dehydrogenase from Ralstonia eutropha
RT "Catalytic and Molecular Properties of the Quinohemoprotein
Strain Bo."
RL U. Bacteriol. 183:1954-1960(2001).
DR EMBL; AF277373; AAF66335.1;
DR HSP; Q924J7; IFLG.
DR InterPro: IPR002372; Bac_PDO_repeat.
DR InterPro: IPR003045; Cyt_c_heme_bind.
DR Pfam: PF01011; Bacterial_PDO_6.
DR Pfam: PF00034; Cytochrome_c_1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SO SEQUENCE 698 AA; 75471 MW; 9C87F180BD1D2FB2 CRC64;

```

Query Match 19.4%; Score 520.5; DB 2; Length 698;
Best Local Similarity 27.4%; Pred. No. 9e-30;
Matches 160; Conservative 80; Mismatches 216; Indels 129; Gaps 21;

```

OY 10 SAGALALAAPFA--QVTPVDELLANPAG--EWISYQNDENRHSFLITQITENYNG 65
Db 14 AAASVALPAPAFGANAAARVDGAIRANEGITPNMPSYGLDAETRFSLQEVNANGVR 73
OY 66 QOLVWARGMPGK-VQVTPLIHDGVNYLANPGDVIOAIDAKTGDLIMHRRQLP----- 119
Db 74 NGLMNSTIDESTGVEATPLVNDGVVYVAPSVVHAIDARKKRLMTYDQVPRDQAY 133
OY 120 -----NIATLNTIVA 128

```

```

Db 134 KCCDDVNVNGVALYOGKVFAGAFDRLVAIDAAGKKVMEODITVDRSKSYTTTGAPRY 193
OY 129 NG-VIVAGSTQCYSPGCFVSGHDSATGEELMRTFIP-----RAGEGDETFMGND 178
Db 194 NGKVIINGCALYVGRG-YITAYIDETGKQGMRYTYVGPAPRPFENEAMAKAAATM--D 250
OY 179 YEAR-WMTG-----AMGQITTPYVNLVHYSTAGPASEFORPTGCTLYGNTREAVP 232
Db 251 PCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
OY 233 PCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
Db 310 PCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
OY 291 GVPCKTGMOPDAETGEELARDTNYOMTIESIDENGIVVNDALTKLELVEYDVCPTL 375
Db 352 HAP-KNGFFVYIDRNGKFLISAKNFVDVWMAAGYDKNGRPVETPOADSGKPA---DVV 406
OY 349 PTEFGDMPGSAALNPDGSIYFPLNNVCYDMMAVDOEFTSMOYNTSNVTKLPKGMIGRID 412
Db 407 PTEFGDMPGSAALNPDGSIYFPLNNVCYDMMAVDOEFTSMOYNTSNVTKLPKGMIGRID 412
OY 398 -----VTKLPKGMIGRIDAIDISTGRITLWSEVERAANYSPVLTGGGVLENGGTDRYF 453
Db 462 LGMLVNAEPPRSKPMGRILAMPDLAOKAVWHRDHAGPMNGGTLAGNLVFGGTADGRLY 521
OY 454 ALSOETGELTNGTRLATYASGOAISYEVDGMOYVAIAGG-----VSYGSLN 501
Db 522 AYHAATGETKMSYETRALYDPEVLTGGDLVFGVIGIDRFDALDAESGKEVWSTRLPGAV 555

RESULT 3
O53362 PRELIMINARY; PRT; 742 AA.
AC O53362;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE
OS Acetobacter pasteurianus (Acetobacter turbidans)
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OC NCBI_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NC11380.
RX MEDLINE-94042848; PubMed-9226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus.
RL U. Bacteriol. 175:6857-6866(1993).
DR EMBL; D13893; BAA40252.1;
DR HSP; Q924J7; IFLG.
DR InterPro: IPR001479; Bac_PDO.
DR InterPro: IPR002372; Bac_PDO_repeat.
DR Pfam: PF01011; Bacterial_PDO_6.
DR PROSITE: PS00363; BACTERIAL_PDO_1; 1.
DR PROSITE: PS00364; BACTERIAL_PDO_2; 1.
SO SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

```

Query Match 18.1%; Score 487; DB 2; Length 742;
Best Local Similarity 26.5%; Pred. No. 2.8e-27;
Matches 161; Conservative 81; Mismatches 215; Indels 150; Gaps 21;

```

OY 11 AGALALAAPFAQVTPVDELL--ANPAGEMWISYQNDENRHSFLITQITENYNG 68
Db 23 AALPVAAPARADQGNWTEGAIITHADHPENWISYQNTSEYGRPLDQINSINNGDLK 82
OY 69 LWARGMPGKQ-VQVTPLIHDGVNYLANPGDVIOAIDAKTGDLIMHRRQLP-NIA----- 122
Db 83 LAMYYTLDITNNGQETPLVNDGVVYVAPSVVHAIDARKKRLMTYDQVPRDQAY 142

```

```

QY 123 --TLN-----I 126
DB 143 CQTVNRGAGYNGKVFEMCTGFDRLVAADATGKKWEVNTIPADSLGKORSYVDGAVR 202
QY 127 VANGVIYAGSTCOYSPFGC--FVSGHDSATGEELMRYNFIIPRAGEGD-----172
DB 203 VAKGLVILGN--GSEFGARGFVSFAFDGKGLKMRFTVPPNKNKEPDHADVNLMSKA 260
QY 173 -ETWMDYEAARMY-----GAMQITTDPTNLYHGSTAVPASSTORCTPGCTLY---223
DB 261 YKTMGP--KGAVYRGGGGYVWDSLVYDPSDLX--LAVG-----NCSPPNNKYRSE 309
QY 224 --GTN-----TRFVAPDGEIYWRHQLPRDNWDOECTFEKAMVTVQVSTEMEGLQSI 277
DB 310 GISNLEFLGSIALKKEGEYVWHFOATPMQMDYTSVQOIMTLDMV-----357
QY 278 NPNAATGE--RVLITGVPCKTGTMQGFDAETGEFLWADPTNYQNMIESIDE--NGIVTVNED 335
DB 358 ----NEMRHVIMHAP--KNGEFTYLDARTGEFLAGKNVYQNMANGLDPLTGRIYVNP 411
QY 336 ALIKELDEYDVCPTEFLGGRDPSAALNPDSCGIFPIPLNNCY-----DMA 382
DB 412 GLYITLGRFVGIPLGPGAHNFMAYSPTKTLVYLPQHIFPGYKNQYGGFKHPDAMN 471
QY 383 VDQFTSMQVNTKLPCKDM-----IGRIDALIDISTGRTLSVERAANYSPVLSTGGGV 442
DB 472 VGLDMTKNGLDPTPE--ARTAYIKDLHGMILLAMPVMEKFTYMKIDHKGPNNGSVLATGDL 530
QY 443 LPNGSTDRYFALSOETGELTMOTRLATVASCALISEYDGMQVYV-----IAGG 492
DB 531 LFQGLANGEFHAYDATNGSDLYKFDQSGIILASPMYISVNGKQYVAVEYGMGIVPISMG 590
QY 493 GYSYSGS 499
DB 591 GVGRTSG 597

```

RESULT 4
Q9EYV8 PRELIMINARY; PRT; 601 AA.

AC Q9EYV8; 092MY9; 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
GN "MAF"
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM1021;
RX Fennel B.J., Tiwari R.P., Dilworth M.J.;
RT "Regulation of C1 assimilation in Sinorhizobium meliloti."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBD databases.
DR EMBL; AF309488; AAC31643.1;
DR HSSP; P38539; 4AAH
DR InterPro; IPR002372; Bac_PQQ-repeat.
DR Pfam; PF01011; Bacterial_PQQ; 7
SQ SEQUENCE 601 AA; 6586 MW; 09F743726B2E2A83 CRC64;

Query Match 17.9%; Score 482; DB 2; Length 601;
Best Local Similarity 25.8%; Pred. No. 4, 6e-27;
Matches 155; Conservative 90; Mismatches 206; Indels 150; Gaps 22;

QY 14 LALLAPAFARQVTPVDEL--LANPAGGEWISYQONDEYHSPSLQITTEWNGQLDLY 71
DB 8 LAIMISGGAGVAFANDELQKLDDP--NOMALQOTGDYANLRKSLQDINKDNVSKLQVAV 66
QY 72 --ARGMQPGKQVYPLIHGCVMTLANP-----GVY-----99

```

DB 67 TESTGVLRGH--EGSPVLIGDLMYVTPPPNTVVALDLSKDQIWMKYEKODPNVLPVNC 125
QY 100 -----IOADAKTGDLIW-----EHRROLPIATINLIYANGV 131
DB 126 CQTVNRGAGYVADNKLIFLQADTTVALADAKTGKTYMSKKNQNDATGECTNATVMPYVDKI 185
QY 132 IVAGSTCOYSPFGC--FVSGHDSATGEELMRYNFI-----PRAGEGD 173
DB 186 LVGISGGEYVRG--HTAYASMABGKVLWGRYSMGPSDYLIDPEKTHLGRVGVKDSGLT 244
QY 174 TWGMDYEAARMY-----AMQITTDPTNLYHGSTAVPASSTORCTPGCTLYGNTTFA 230
DB 245 TWEGD--QMKIGGTTWGYSTDEPNELVYGTGNPTWNPQTR--FGDNRW--SMTIFA 298
QY 231 VRPDTGLWRHQLPRDNWDOECTFEKAMVTVQVSTEMEGLQSIIPNATGERRVLT 290
DB 299 RDVDTGMAKMLYQMTPRHDEWDYGVNEMILITGQHLIDGR-----DRKLLT 342
QY 291 GVPCKTGMQFDETEGFL-----WARDTVYQNMIE-----SIDENGI 329
DB 343 HED--RNGFGYTMDEVTELLVAEKYDPTVWATVEYVMDKSKYGRPVYVAYSTEDNG- 400
QY 330 VIVNEDALIKELDEYDVCPTEFLGGRDPSAALNPDSCGIFPIPLNNCYDMAVDOEFTS 389
DB 401 ----EDTNT-----GCPALGKTKDOOPAAVSPKTELYVPTNHCADYEFPRSYTA 450
QY 390 MDVYNTSNVTKLPCKDM-----IGRIDALIDISTGRTLSVERAANYSPVLSTGGGV 446
DB 451 GQPVGATLSMPY--KDSHGGMNFIACDNKEGKIKMSLPESFVSGALATAGDVVFTG 509
QY 447 GMDRYFALSOETGELTMOTRLATVASCALISEYDGMQVYV-----GYSYSGSLN 501
DB 510 TLEGIKAVDAATGKELRYFRKTPSGVIGVMTYAREGQYAVVLSGVGNAGIGLALGTL 569
QY 502 S 502
DB 570 N 570

```

RESULT 5
Q9ZMY9 PRELIMINARY; PRT; 601 AA.

AC Q9ZMY9; 092MY9; 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
GN SMD20173.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RX Finlan T.M., Weldner S., Wong K., Bumester J., Chain P., Gouzy J.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puchler A.,
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AF603642; CAC48573.1;
DR Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 601 AA; 65759 MW; D73424EFBD15ADBE CRC64;

Query Match 17.9%; Score 481; DB 16; Length 601;
Best Local Similarity 25.8%; Pred. No. 5, 7e-27;
Matches 155; Conservative 90; Mismatches 206; Indels 150; Gaps 22;

```

OY 14 LALLAPAFQVPTVDEL--LAMPAGEMISTGONENYRHSPLQITENYGOLOLW 71
DB 8 LAIMSIGGAQVAFANDELQIDLP--NOMAIOTGVANLRYSKLOIKNDVNGKLQVAM 66
OY 72 --ARMOGKQVPTPLHGDVMTLNDP-----GDV----- 99
DB 67 TFSIGVLRGH--EGSPVLYGIDLMVYTFEPNTVYALDLSGQVYKPEKODPNVYPMC 125
OY 100 -----IOAIDAKTGDLT-----BHRQPLNATLNVANGV 131
DB 126 CDVNRGVAYADNKFILHQADTIVYALDQKATGKIVMSYKNGDATGNTATVMPVKDKI 185
OY 132 IVASGTQYSPFGCFVSGHDSATGEELMRYFT-----PRAGEGDE 173
DB 186 LVGSGGEFVGRG--HYTAISMADGKVLNMGYSMGPSDPLLDPEKTHLGRPVGKSGLT 244
OY 174 TWGNDYFARMWNG--AMQITVDPVTLNVHGSFVAVPASSETORGPBGGLYGTNTRFA 230
DB 245 TWBGD---QWKIGGGTGWMSYDEPENLVYGTGNSTMPNTPOR--PGDNRM-SMTIFA 298
OY 231 VRDPTGEIYWRHQTLPROMWDECTFEEMVTVNDVOSTEMEGLOSINPMAATGERVLT 290
DB 299 RQDPTMAKWLQMTPEHMDVDGVNEMILEQIDCK-----DKLLT 342
OY 291 GVPCKTGMQDPAETGEFL-----WARDTNQNMIE-----SIDENGI 329
DB 343 HFD--RNGGCTMDRVYTGELVAEKYDPTVKNATELVMDPSDKYGRPOVVAQSTEQNG- 400
OY 330 VTNEDALIKELDEYDVCPTFLGGRWPSALNPDSGIYFIPLVNVCYDMAVQDEFTS 389
DB 401 ----EDWTIT-----GVCPSALGTGKDOCPAAYSPTKELTFYPTNVNCHDDEPRVSTYA 450
OY 390 MDVYNTSVNTKLPKDKM--IGRIDAIDISTGRTLMSYERAAANTSPVLTSGGVLENG 446
DB 451 GQPVGATLSYPP--KDSHGMGNFAMTAMNKEKIKMSLPEPVSWSALNAGVYVYFG 509
OY 447 GTDYPFALSOEGETIMOTRLTFVASGSAISTEVDGMQVYIAGI-----GVSGSLN 501
DB 510 TLEGYLAAVDAATGKELRYKTPSVIGVGNMTYAREGKQVAVLGVGAGIGLACLT 569
OY 502 S 502
DB 570 N 570

RESULT 6
ID 046444 PRELIMINARY: PRT: 708 AA.
AC 046444:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE QUINOHAEOMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
DE (EC 1.1.99.-) (OH-EDH1).
GN OHEDH.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID-285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15667;
RX MEDLINE-96184549; PubMed-8654419;
RA Stoerovogel J., Kraayveld D.E., van Suijs C.A., Jongejan J.A.,
RT "Characterization of the gene encoding quinohaemoprotein ethanol
RL dehydrogenase of Comamonas testosteroni.";
RN Eur. J. Biochem. 235:690-698(1996).
RP SEQUENCE OF 32-54 AND 477-490.
RC STRAIN-ATCC 15667;
RX MEDLINE-95324580; PubMed-7601151;
RA De Jong G.A.H., Geertlof A., Stoerovogel J., Jongejan J.A., De Vries S.,
Dune J.A.,

```

"Quinohaemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purification, characterization, and reconstruction of the apoenzyme with pyroloquinoline quinone analogues."; Eur. J. Biochem. 230:899-905(1996).

RP CHARACTERIZATION.
RX MEDLINE-86242113; PubMed-3521592;
RA Groen B.W., van Kleef M.A., Dune J.A.,
RT "Quinohaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
RL testosteroni.";
RN Biochem. J. 234:611-615(1986).
[4]

RP CRYSTALLIZATION.
RX MEDLINE-21536088; PubMed-11679760;
RA Oubrie A., Huizinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
RT Dune J.A., Dijkstra B.W.;
RL "Crystallization of quinohaemoprotein alcohol dehydrogenase from
Comamonas testosteroni: crystals with unique optical properties.";
CC Acta Crystallogr. D 57:1732-1734(2001).
CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC -1- CRYSTALLOGRAPHIC SPACE GROUP: P2₁/2₁/2₁.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBUNIT LOCATION: PERIPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
DR EMBL; X81880; CA57464.1; -.
DR HSP: 092477; 1PUG.
DR InterPro; IPR002372; Bac_PQQ_repeat.
DR InterPro; IPR003088; Cyl_C1.
DR Pfam; PF01011; Bacterial_PQQ_C1C.
DR Pfam; PF00034; Cytochrome_c_1.
DR PRINTS; PR00605; CYTOCHROME_C1.
DR SIGNAL; PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
KW SIGNAL 1 31
FT CHAIN 32 708
FT BINDING 635 635 QUINOHAEOMOPROTEIN ETHANOL DEHYDROGENASE
FT BINDING 638 638 TYPE 1.
FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
SQ SEQUENCE 708 AA; 76822 MW; 99AB54BD66ACAB3 CMC64; (BY SIMILARITY).

Query Match 17.8%; Score 479.5; DB 2; Length 708;
Best Local Similarity 25.0%; Pred. No. 9.3e-27;
Matches 148; Conservative 84; Mismatches 238; Indels 121; Gaps 18;

```

OY 3 PTLNMSAGALALLAPAFQVPTV-----DELLANP-AGEMISTGONEN 50
DB 12 PGRMWLILAAAG--SAAFAQTGPAQAQAAVAVRVDGDFIRANAARTPMPPIGVDAE 69
OY 51 YRHSPLQITENYGOLOLWARKMQPEK-VQVPTPLHGDVMTLNDPQIDAKTGD 109
DB 70 TRSRIDQINAAVADGLSLAWSYMLSTREYVATPVVDCIMYVSSMSVVAIDTRGN 129
OY 110 LTWEHRQLPNT--ATLVANG----- 130
DB 130 RIMTYDPOIDSTGFGKGCDDVNVNGVALMKGKYVGAAMDRLAIDATGKEWHQNTFE 189
OY 131 -----VVASGTQYSPFGCFVSGHDSATGEELMRYFT----- 165
DB 190 GOKSLITGAPRVKFKVITIGRGALEYGRG-YITAYDAETGKRWKFVSPGDPSPKPF 248
OY 166 --RAGEGDETWGNDYFARMWNGA-----NGQITVDPVTLNVHGSFVAVPASSETORGP 218
DB 249 EDESMKRAARTW--DPSGKMWEGGGCTWMDSMITDAELNITVYGTGSGFWSHRYVSPK 306
OY 219 GGLTYGNTREFAVRDPTGEIYWRHQTLPROMWDECTFEEMVTVNDVOSTEMEGLOSIN 278
DB 307 GGNILYASIVLADPDTGKYKWHYOTPGDMDYTSQPMILADIKI----- 353

```


RESULT	8	PRELIMINARY:	PRT:	629 AA.
Q9AC48	09AC48			
ID	09AC48	PRELIMINARY:	PRT:	629 AA.
DC	09AC48			
DT	01-JUN-2001 (TREMBLrel, 17, last sequence update)			
DT	01-JUN-2001 (TREMBLrel, 17, last sequence update)			
DT	01-DEC-2001 (TREMBLrel, 19, last annotation update)			
DE	METHANOL DEHYDROGENASE MXAF.			
GN	MXAF.			
OC	Methylobacterium nodulans.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Methylobacterium group; Methylobacterium.			
NCBI_TaxId	114616;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-OR52060;			
RC	MDLNA-20566686; PubMed-11114919;			
RX	SY, Giraud E., Jourdan P., Garcia N., Willems A., de laJaudie P.,			
RA	Pfin Y., Neyra M., Gillis M., Bolvin-Masson C., Dreyfus B.			
RT	"Methylotrophic Methylobacterium bacteria nodulate and fix nitrogen in			
RT	symbiosis with legumes."			
RL	U. Bacteriol. 183:214-220(2001).			
EMBL	AF220764; AAC49450.1; -			
DR	HSSP; P38539; AAAH.			
DR	InterPro; IPR001479; Bac_POO.			
DR	InterPro; IPR002372; Bac_POO_repeat.			
DR	Pfam; PF01011; Bacterial_POO_1; 1.			
DR	PROSITE; PS00363; BACTERIAL_POO_2; 1.			
DR	PROSITE; PS00364; BACTERIAL_POO_2; 1.			
DR	SEQUENCE 629 AA; 68798 MW; 3DB559DB373BB0B2 CRC64;			

QY 173 -----ETWGDYEARMTGA---WGOITYPVTLVHGSTAVGPASE 212
 Db 242 KQENIHNAHYGOKGLSTWEGD---AMKIGGTMGMWYADPGNLTITGTPAPWME 298
 QY 213 TORPGGTLYGTTRAVRPDPTGELVNRHQTLPDRNDQDCTEEMAVTVNDVOPSTEME 272
 Db 259 TMR--PDGNK--TWTFIARVDVTGEAKFCYOKTPHDEMDIAGVNMML-----STOKD 348
 QY 273 GLOGINNAATGERRVLTGVPCKTGMOPDAETGEFLMAR--DNYNMIESIDENGI 329
 Db 349 -----RSGKERLLTHPDNRNGIYTTLDRTNGDLSAHKIDT--VNEFKYDLKSG 397
 QY 330 VVNEDEALIKEID-VEYVYCEPTFGGNDPSPALNPDGTYFPLNNVCYDMAVDOEFTS 388
 Db 398 LPRVDEYGTBRDHLAKICPSAMGYHNOCHSDYDEPERLTFYMGINHICMDEPFMLPYR 457
 QY 389 SMQVYTSNVTKLPPGK-----DMIGRIDAIDISTGRTLSVERAANYSPLYSTGGV 442
 Db 458 AGQFVGAATLMMYGPFGKGDROMAEGQIKAYDALTGKFKWKKERRAVMGGLTATAGVY 517
 QY 443 LFNGSTDYFRALSQETGETLMQTRLATVASGAISTYVDGMQYVAIAGGVSYSYG 499
 Db 518 VFYGLDGFIKARHSDTGLMMAKLPSCALGYEPTVTHKGTQYAL-----YGVG 569
 RESULT 9
 P71509 PRELIMINARY: PRT: 601 AA.
 AC P71509;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MKAP.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OX Methylobacterium group; Methylobacterium.
 RN NCBI_TaxID=408;
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=97312011; PubMed=9168622;
 RT Chistoserdova L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RL methyloctrophy gene clusters in Methylobacterium extorquens AM1";
 DR EMBL: U72662; AAB58890.1;
 DR HSSP: P38539; AAAH.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 SO SEQUENCE 601 AA; 64952 MW; 68E45C7059CBB239 CRC64;

Query Match 17.2%; Score 463.5; DB 2; Length 601;
 Best Local Similarity 25.4%; Pred. No. 1.e-25;
 Matches 152; Conservative 97; Mismatches 220; Indels 129; Gaps 22;

QY 1 MKPTSLMASAGALALLAPARAQVTVYTDLLANPAGEMISGXNOENYHNSPLQIT 60
 Db 1 MRAVHLALGAGLAA--ASPALANESVYLG--YANP--AAQVLOTVDYANTRYSLDIN 54
 QY 61 TEWVGLOLVW--ANGMOPGKVOYVPLIHGVNVLNP----- 96
 Db 55 ASNVKMLQVAMTFSTGVLRGH--EGSPLYGNIMYVTPPPNIVVALDIDOGAKIVMYEP 113
 QY 97 -----GDVIOAIDAKTGDILMEHRQLP-----N 120
 Db 114 KODPSVIFVMCDDIVNRGLAVADGAILLHQADTLTVLSIDAKSGKNNYSVKNGBDSKGEIN 173
 QY 121 INTLNIIVANGVYVAGTCQYSPFGCVSGHDSATGEELMRYET----- 164
 Db 174 TATVLPVKDKAVYVIGISGEGF--VOCHVYAYDLKSGKVMKSTISGDDQDLIVPEKTSLS 232

QY 165 --PRAGEGDETMGNDYEARMTG---WGOITYPVTLVHGSTAVGPASE 212
 Db 233 GRPIAKDSLSKTWEGD---OMKGGGCTGWGSPKIDLWYSG--NSTNPNKQRBG 287
 QY 220 GLVGTNTRAVRPDPTGELVNRHQTLPDRNDQDCTEEMAVTVNDVOPSTEME 272
 Db 259 TMR--PDGNK--TWTFIARVDVTGEAKFCYOKTPHDEMDIAGVNMML-----STOKD 348
 QY 273 GLOGINNAATGERRVLTGVPCKTGMOPDAETGEFLMAR--DNYNMIESIDENGI 329
 Db 349 -----RSGKERLLTHPDNRNGIYTTLDRTNGDLSAHKIDT--VNEFKYDLKSG 397
 QY 330 VVNEDEALIKEID-VEYVYCEPTFGGNDPSPALNPDGTYFPLNNVCYDMAVDOEFTS 388
 Db 398 LPRVDEYGTBRDHLAKICPSAMGYHNOCHSDYDEPERLTFYMGINHICMDEPFMLPYR 457
 QY 389 SMQVYTSNVTKLPPGK-----DMIGRIDAIDISTGRTLSVERAANYSPLYSTGGV 442
 Db 458 AGQFVGAATLMMYGPFGKGDROMAEGQIKAYDALTGKFKWKKERRAVMGGLTATAGVY 517
 QY 443 LFNGSTDYFRALSQETGETLMQTRLATVASGAISTYVDGMQYVAIAGGVSYSYG 499
 Db 518 VFYGLDGFIKARHSDTGLMMAKLPSCALGYEPTVTHKGTQYAL-----YGVG 569
 RESULT 10
 O9AF95 PRELIMINARY: PRT: 691 AA.
 AC O9AF95;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OX Pseudomonas.
 RN NCBI_TaxID=86174;
 RP SEQUENCE FROM N.A.
 RC Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT "Characterization of the expression of two distinct alcohol
 RL butanovora";
 DR EMBL: AF355798; AAK27220.2; to the EMBL/GenBank/DBJ databases.
 SO SEQUENCE 691 AA; 75070 MW; 4FC7ED20CDA14E64 CRC64;

Query Match 17.2%; Score 461.5; DB 2; Length 691;
 Best Local Similarity 26.4%; Pred. No. 1.9e-25;
 Matches 145; Conservative 82; Mismatches 205; Indels 117; Gaps 19;

QY 39 GEMWISYGOENYHNSPLQITTEENYGOQLYVARGMOPGK-VQVPLIHGVNVLNP 100
 Db 41 GEMRHGIDYDAGTRYSPLAQITPDNAKELGLVMSYDLESSRGVEARPIVYDGMVYAPW 100
 QY 98 DVIOAIDAKTGDILMEHRQLP-----NATLNIIVANGVY-ASTQYSPFGCVSGHDS 152
 Db 101 SVVHALDVRSGRKMLTYDDEYPEREKNAACDYYNRYAVNAHEKRYVGSGLDRIVA-IDA 159
 QY 153 ATGEEELM-RNYET-----PRA--GEDEDTWGNDEYARMTGTAM-----SQ 190
 Db 160 RTGKRWERNTLIDDKPYTTGAPRVYIKKVVYVINGGAEEFVGRYITAYDPTAASRPV 219
 QY 191 I-----TYDPTNLVHGSTAVG----- 208
 Db 220 VPGDDPSLPEDASMAAKATWDPAGVYVYVINGGRRHGVLELDGLYKAGCCYSAPATPS 279
 QY 209 PASFQRTGCTPGTLYGNTFAVRPDPTEIYVNRHQTLPDRNDQDCTEEMAVTVNDVOPSTEME 272
 Db 280 PWSHKKSPAGGDLXTASIVALRPDTEIYVNRHQTLPDRNDQDCTEEMAVTVNDVOPSTEME 272
 QY 267 PSTMEGLQSIINNAATGERRVLTGVPCKTGMOPDAETGEFLMARNTNYNMIESIDE 326

Fri May 24 11:27:28 2002

walick-934-125.pcp.rsp

Page 7

```
Db 340 P-----RKVILHAB-KNGFEFFIDRDKFKITSQNPVNNMAGTDE 380
QY 327 NGIATVN-EDALIKELDYEDVCEPTFLGGRDMSFSAALNPDGTYIFPLNNVCYDAMAAVQ 385
QY 381 NGRIENPGAMPGLHSMRF---PAPSARINWHSMSYSPQGLAYFPANIPL-VLQEDK 436
QY 386 EFTSMQVNTSN-----YTKLPKDKMIGRIDALIDISTGRTLMSVERAAA 430
Db 437 NMS-----YNOAQPGQAMAGIGMNLGMLNPRPASQPGRLIANDPVQOKVEVRKEHVSP 492
QY 431 NISPVSTGGVLENGGDRYFRALSOETGETLMOTRLATVYASQALSYEDVMQVYAL 490
Db 493 WNGTGLTAVGNVFGSTADARILALDARDGKELMSAPMGTVIAPVYEVDRKQVYISIA 552
QY 491 -GGGVSYGS 498
Db 553 VGMGVYGN 561

RESULT 11
ID 024759 PRELIMINARY; PRT: 633 AA.
AC 024759;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE METANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN MXAR.
OS Hypnomicrobium methylotrophum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypnomicrobium group; Hypnomicrobium.
OX NCBI_Taxid=84;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM2;
RC MEDLINE-97457202; PubMed-9311140;
RX Tanaka Y., Yoshida T., Matanabe K., Izumi Y., Mitsunaga T.;
RA "Cloning and analysis of methanol oxidation genes in the methylotroph
RT Hypnomicrobium methylotrophum GM2."
RL FEBS Microbiol. Lett. 154:397-401(1997).
DR EMBL; AB004097; BAA2372.1; -.
DR HSSP; P38539; 4AAB.
DR InterPro: IPR001479; Bac_PQO.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR Pfam; PF01011; Bacterial_PQO_7.
DR PROSITE; PS00363; BACTERIAL_PQO_1; 1.
DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
DR Signal; Oxidoreductase.
KW SIGNAL.
FT CHAIN 35 633 POTENTIAL.
FT SIGNAL 1 34
FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;
SQ

Query Match 16.0%; Score 431.5; DB 2; Length 633;
Best Local Similarity 24.1%; Pred. No. 2.6e-23;
Matches 146; Conservative 96; Mismatches 214; Indels 151; Gaps 22;
```

```
QY 165 -----PRAGEE--GDETWNDEYARMMGA---WGQITDPVNTLVHGS 204
Db 237 PDEILADDFNSANPHYGOKGLGTATWEGD---AMKIGGTNMGWATYDPAQNLIIYGS 293
QY 205 TAVGPASETQRCRGPGTLYGTNTFRAPVPTDGTGLVWRHQTLPENNMDECTEFEMMYTND 264
Db 294 GNPAPMETWR--PGDNK--TWITARADDTGKMKFGYOKTPHEMDFAGVNVIMLSE-- 348
QY 265 VQSTEMEGLOSINPNAATGERVLTGVPCKTGTWMOFDAETGEFLMA--RDTNYONMI 321
Db 349 ---QDKE-----GKRRKLLTHPDRNGIYITLDRENGDLISADKDDT--VNVF 392
QY 322 ESDENGIVYNEDALIKELDVE-YDVCEPTFLGGRDMSFSAALNPDGTYIFPLNNVCYD 380
Db 393 KHVDSKGLPYRDEPFGTRMDHKGTCEIPSAMGYHNGHDSYDPTKOLFPGMINKICMDW 452
QY 381 MAVDQETSMQVNTSNVTKLP--PGKDM-----IGRIDALIDISTGRTLMSVERAAANSP 434
Db 453 EEFMLPYRAGQFVGATLMTYPGKGRONYLIGQIKAYNAITNTYKHEMERFSVWG 512
QY 435 VLTGGGVLENGGDRYFRALSOETGETLMOTRLATVYASQALSYEDVMQVYALAG-G 493
Db 513 TLATAGNLVFTYITDGLFLKARNSDTGELMKRLPSGVIGTPMTEHKVQYIATVMSYG 572
QY 494 VSYGSL 500
Db 573 GMPGYL 579
```

```
RESULT 12
ID 09AGW3 PRELIMINARY; PRT: 623 AA.
AC 09AGW3;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (DPC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1; -.
DR HSSP; Q924J7; IFUG.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR Pfam; PF01011; Bacterial_PQO_4.
KW SIGNAL.
FT CHAIN 1 29 POTENTIAL.
FT SIGNAL 1 29
FT SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;
SQ

Query Match 15.9%; Score 428; DB 2; Length 623;
Best Local Similarity 24.9%; Pred. No. 4.6e-23;
Matches 157; Conservative 81; Mismatches 217; Indels 176; Gaps 24;
```

```

OY 164 -----IPRAGEGDDTM-----175
Db 181 YTMGAPVLIKDKSGRTLLVHGSSGDEFGVGNLFARDPDTGEEVWARRPVEGSHMGRLN 240
OY 176 GND-----YEAKMTGA---WGQITDPYTNLYHGSTAVGA 210
Db 241 GKSTPTGDKPAPSWPDDNSPTGKVEA-WSGGGGAPQOTASFVENNMVYIGAGNAPW 299
OY 211 SEDORGAPG-----GTLCYNTRAVBPDTGEIYWRHOTLPBNDQDETFEMATYN-- 262
Db 300 NTKKRAAPGDDPRNDMSLF-TSGQAYVDASTGELKGFYQHTPRNDAMDFSGNSNVLEEK 358
OY 263 -----VDOPSTMEGJOSIN-----PNAATGEBRYLVGPKCTGTMMQ--F 302
Db 359 DPKGTMMVNASAHADRNGEFFYTRDMLAKAGAPYKPTSL-LIGAMFVVGITWASGF 415
OY 303 DAEQTGFLMARPTYNQNNIESIDENGIVYVNEALIKELDEYDVCPTLGGSDMPASAL 362
Db 416 DIKTKRPL-EKDNRRPPOKREGADGESIFVS-----PPLGSTNNHMSY 459
OY 363 NPDSSIVYPLNNVCYDMMAVDEFTSMDYXNTSNT-----KLPPG-----KDMIG 409
Db 460 SPDTGLFFIIPANHAMD-----IWTENYTYAGSAUYGOGFRIKLNLEDDHYG 506
OY 410 RIDMIDISTGTLMSVERAANTSPVLSTGGVLENGSTDRYFRALSOETGETLMOTRLA 469
Db 507 ILKRIDSPPARISGAOGRFPA-VAGTLTYAGSMVFPTGSDGLKAFDAKNGKELMKFPOTG 565
OY 470 TVASGAISYEVDMQIYVAIAGGVSYGGL 500
Db 566 SGVVSVYVTEWMDGEGYVALIQSG---YGVAV 593

```

RESULT	13		
ID	09L935		
AC	09L935	PRELIMINARY:	PRT; 599 AA.
DT	01-OCT-2000	(TREMBLrel. 15, Created)	
DT	01-OCT-2000	(TREMBLrel. 15, last sequence update)	
DE	METHANOL DEHYDROGENASE LARGE SUBUNIT.		
GN	MXAF.		
OS	Methylovorus sp. (Strain SSI / DSM 11726).		
OC	Bacteria; Proteobacteria; beta subdivision; Methylophilus group;		
OX	Methylovorus.		
NCBI	taxid=81683;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SSI;		
RA	Kim Y.M.;		
RT	"Cloning and nucleotide sequence of mxaf gene of Methylovorus sp		
RT	strain SSI DSM11726."		
RL	Submitted (SEP-1999) to the EMBL/genbank/DBJ databases.		
DR	EMBL; AF184915; AAD56237.2; -		
DR	HSSP; P38539; 4AAH.		
DR	InterPro; IPR001479; Bac_POO.		
DR	InterPro; IPR002372; Bac_POO_repeat.		
DR	Pfam; PF01011; Bacterial_POO; 7.		
DR	PROSITE; PS00364; BACTERIAL_POO_2; 1.		
SQ	SEQUENCE 599 AA; 6513 MW; DBF6F4B5D871BC91 CRC64;		

Query Match	15.68;	Score 419;	DB 2;	Length 599;
Best Local Similarity	24.48;	Pred. No. 2e-22;		
Matches 142;	Conservative	ac		

[illegible]

```

Db      64 YKNLSSGIGHSTELLIGH-EGAPVIGDMVYINTEPPNNFALNLAPREKIVQHKPKO 1222
OY      116 -RQLENTATLNIYANGV-----IVA-----GSTCOY 140
Db      123 DASVKAACCDIYRGLAYDDGHIFFKTQDGLHLVASDAGTKELMKMKNCPAVGSTITQ 182
OY      141 SPF-----GC-----FVSHDSATGSEELMKNRYFT-----PRA 167
Db      183 APEAVRGLVLYGSGAELEGRVYATVYDQKGTGELVWRSFATGPBNELNLAKFNKNPHY 242
OY      168 GEE--GDEITGNDYEAARMGTGA--WQOITDPVYINLYHGSTPVGASSTOGTGTGTL 222
Db      243 GQGLGKTGWEEB---GMKIGGGTNMGWYAVDPKLTNIFYGSGNPAPMNETM--PGDMK 297
OY      223 YGNTFPFARPDPTGYVWHOTLPDPMDNDQCTFEMAVTVNDVOPSTEMEGLOSINPNA 282
Db      298 W-TMTIWARDVDTGAANKGYOKTPhDENDEPRAGVNMOLITDOAVNGKTO----- 344
OY      283 TGBRRLLVGPCKSTIMQOPAEITGEFLMARDTN-YQNMIESIDENGIYVNEADILKEI 341
Db      345 -----PLTHVD-RNGIMYITLNRQTSIYOAAKVDAVAVNFKKVLDLKTGLPRHDEESTRM 399
OY      342 DVE-YDVCPTEFLGGRDWPSPALANDSDSITYPIPLNNCYDMAAVDOEFTSMDYNTSNYTK 400
Db      400 DHRGTNIDPSAMGEFHNGDLAYDPDSTITYFTGLNHCMDMEBFMLPYRAGOFYFGATITAM 459
OY      401 LP-----PCMDMIGLIDAIDISTORTLMSYERAANYSPVLSTGGGVLEPNGSTDYFPALS 456
Db      460 YPGNGPPTKREMGVYGLAMDGVTEGVKWTMEKFSYWGGLTLATKGGLVFPYNTLDGNITKALD 519
OY      457 QETSEITLMQRLATVASGAISIEYVDMQYVAILAGGVSYSSG 499
Db      520 KTNKGELMKFKMPSGAIGAPMSVAYKMGQYLA-----TNGVG 557

```

```

RESULT 14
0934G0
ID      0934G0
AC      0934G0;
          PRELIMINARY;
          PRT: 695 AA.
DT      01-DEC-2001 (TREMblrel. 19, Created)
DT      01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      LUPANINE HYDROXYLASE PRECURSOR.
OS      LUP.
OC      Pseudomonas sp. DH2001.
OC      Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX      Pseudomonas.
OX      NCBI_TaxID=163360;
          [1]
RA      HOPPER FROM N.A.
RA      Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
RA      Roczinski J.;
          "Cloning, sequencing and analysis of the gene for lupanine
          hydroxylase, a quinoxichrome c from a Pseudomonas sp."
          Submitted (JUN-2001) to the EMBL/Genbank/DBD databases.
          EMBL: AJ318095; CAC67410.1.
KW      Signal.
FT      SIGNAL.
FT      CHAIN.
FT      1
FT      26
FT      27
FT      695
FT      695
FT      LUPANINE HYDROXYLASE.
FT      SEQUENCE 695 AA; 74956 MW; 0996C6387E457FE0 CRC64;

```

Query Match	15.0%;	Score 404.5;	DB 2;	Length 695;
Best Local Similarity	25.3%;	Pred. No. 2.8e-21;		
Matches	136;	Conservative	73.	Matches

OY		38	A G E I W S Y G A O N E V R H S P L T Q I T T E V G C D L M A B G K Q P K R Y O V - P L I H D E V M T I A N P	96
D b	:	36	S G M S I L L G G E N E D H Y S A L K D Y N K S M X N L I G S F T D M E A G D G L V A G P L A D V I I R Q G P	95
OY		97	G D V I A L D A T G G L I E H R R O L - P N I A T I N I A N G V I A S G T C O - - - -	139
D b	:	96	P G K I Y A N D L E T G K N I L T T P P E Y O X D K I T S M T G F W G T H V N G L A V D D N N Y I S S Y C L L A V	155

Fri May 24 11:27:28 2002

walick-934-125.pep.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:18 : Search time 101.42 Seconds
(Without alignments)
553.069 Million cell updates/sec

Title: WALICK-934-128.PDP
Perfect score: 2705
Sequence: 1 MKPTSLIMASAGALALIAAP.....GMQVYALIGGSVSGSLINS 505

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2669.5	98.7	579	AAW37873	Alcohol and/or ald
2	2344.5	86.7	579	AAW37874	Alcohol and/or ald
3	2310	85.4	579	AAW37876	Alcohol and/or ald
4	2251	83.2	578	AAW37875	Alcohol and/or ald
5	480.5	17.8	754	AAW35987	Sorbitol dehydrog
6	479	17.7	738	AAW35987	ADH complex prote
7	478	17.6	738	AAW35987	ADH complex prote
8	476.5	17.6	742	AAW35987	Amino acid sequenc
9	314	11.6	740	AAW35019	Sorbitol dehydrog
10	155.5	5.7	443	ABG24430	Novel human diagno
11	134.5	5.0	824	AAW09614	Pathogenic Staphyl

12	134	5.0	715	22	AAW81738
13	133.5	4.9	910	22	AAW83007
14	133	4.9	966	22	AAW82914
15	133	4.9	1016	22	AAW34349
16	130	4.8	470	22	AAW82701
17	129	4.8	948	22	ABG25990
18	129	4.8	1510	22	ABG21573
19	117	4.3	593	11	AAW07014
20	116	4.3	480	8	AAW70468
21	115.5	4.2	594	12	AAW10005
22	113	4.2	404	22	AAW82772
23	110	4.1	269	11	AAW07006
24	110	4.1	269	15	AAW83295
25	109	4.0	269	10	AAW94788
26	109	4.0	593	15	AAW62944
27	108.5	4.0	530	22	AAW12236
28	107.5	4.0	886	22	AAW33996
29	107	4.0	501	18	AAW10645
30	106.5	3.9	2599	21	AAW75098
31	106	3.9	501	18	AAW10644
32	106	3.9	514	22	AAW85225
33	106	3.9	770	22	AAW30824
34	106	3.9	966	19	AAW49077
35	105.5	3.9	429	22	AAW6542
36	105.5	3.9	469	21	AAW85578
37	105.5	3.9	278	22	AAW82625
38	105	3.9	501	18	AAW26108
39	105	3.9	501	18	AAW26107
40	105	3.9	501	18	AAW26107
41	104	3.8	501	18	AAW26107
42	103.5	3.8	493	22	AAW33026
43	103.5	3.8	649	22	AAW65538
44	103.5	3.8	974	19	AAW49079
45	103.5	3.8	1376	22	ABW52592

ALIGNMENTS

RESULT 1	AAW37873	standard; Protein; 579 AA.
ID	AAW37873	
XX	AAW37873	
AC	10-AUG-1998	(first entry)
XX		
DE	Alcohol and/or aldehyde dehydrogenase A	amino acid sequence.
XX		
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulonate; L-ascorbic; inhibition.	
OS	Glucanobacter oxydans.	
XX		
XX	Key	Location/Qualifiers
FT	Peptide	1..23
FT		/note="signal peptide"
FT	Protein	24..579
FT		/note="mature protein"
XX		
PN	EP832974-A2.	
XX		
PD	01-APR-1998.	
XX		
PF	11-SEP-1997;	97EP-0115801.
XX		
PR	19-SEP-1996;	96EP-0115001.
XX		
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
XX		

S. epidermidis ope
S. epidermidis ope
S. epidermidis ope
Staphylococcus aur
S. epidermidis ope
Novel human diagno
Protein G variant
Sequence of polype
Streptococcus GX78
S. epidermidis ope
Protein G variant.
IgG-binding Strept
Protein G variant.
Streptococcus prot
Human PRO340 poly
Human protein sequ
Human protein L1 pro
HPV6 mutant L1 pro
Neisseria meningit
HPV6 mutant L1 pro
Human protein sequ
Amino acid sequenc
Solium tuberosum
Human interferon-a
Human protein sequ
Human prostate can
S. epidermidis ope
HPV11 mutant L1 pr
HPV11 mutant L1 pr
C glutamicum prote
Drosophila melanog
Solium tuberosum
Escherichia coli p

DR WPI: 1998-195228/18.
 DR N-PSDB; AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA;

Query Match 98.7%; Score 2669.5; DB 19; Length 579;
 Best Local Similarity 90.8%; Pred. No. 2,3e-227;
 Matches 505; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

OY 1 MKPTSLIMASAGALALAAPAAQVTPYTDLLANPAGEWISYQONQENRHSPLTQIT 60
 DB 1 mkptslimasagalalAAPAAQVTPYTDLLANPAGEWISYQONQENRHSPLTQIT 60
 OY 61 TENVGQLOLVWARGKQKQVOTPLIHGQVWYLANPGVIOAIDAKTGDLIWEHRQLPN 120
 DB 61 tenvgqlqlvwarqmgpkqvotplihdgvywylanpgdivaiddaktdglwehrrqlpn 120
 OY 121 IATLNSFG-----I 129
 DB 121 iatlnsfgeptrgmalylgtvnyfsvwdnhlvaldtatgvtfdvdrqggedmvsnsqpi 180
 OY 130 VANGYVAGSTCOYSPFCEVSGHDSATGEBELNRNFIFRAGEGDEMGNGNYEARMTG 189
 DB 130 vangyivagstcoyspfcevsghdsatgelnrnfifragegdegmngnyearmtg 189
 OY 181 vanyvlgvsgtscqspgctvsgndsatgeelwnyflpraseegdetvngdyearmtg 240
 DB 181 vanyvlgvsgtscqspgctvsgndsatgeelwnyflpraseegdetvngdyearmtg 240
 OY 190 AMGOITTPVNTLVHGSTAVGASFTQGTGEGTLYGTNTRFARVPDTEIWRHQLP 249
 DB 241 awgiltvdpvntlvhgstavgasetqgtgtygtlgntrfavrptdteiwrrhqlp 300
 OY 250 RDNWDOCTFEKMTVNDVQVSTFEMGLQSIINPNAATGERRYLGVPCKTGTMOQDAET 309
 DB 301 rdnwdogctfemmtvndvqvpstfemglqsinpnaatgerrrylgvpccktgtmwdqdaet 360
 OY 310 GEFLLARDTYONMIESIDENGIVVNEBALIKELDVEYDVCPTFLGGRDWPSSALNPDS 369
 DB 361 gefllardtyonmiesidengivvnedalikeidveydvcptflggrdwpssalnpds 420
 OY 370 GIFIFLANNVCYDMAVDOETSMQVNTSNVTKLPPKMKMIGIDIDISTGTMTASVE 429
 DB 421 giyiflplnnvcydmavdofetmsmdvntsnvtklppkmmigiddidistgtmtasve 480
 OY 430 RAAANSPVLSITGGVLTNGSTDRFRALSQETGETIMQTRLATVASQAISIEVNGMOY 489
 DB 481 raaanyspvlstggvltngstdrfralsqetgetimqtrlatvasgaaisievngmoy 540
 OY 490 VAIAGGVSYSGSLNS 505
 DB 541 valaggyvsgysgslns 556

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein; 579 AA.
 XX
 AC AAM37874;
 XX

DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..23
 FT Protein /note="signal peptide"
 FT Protein 24..579
 XX /note="mature protein"
 XX
 PD EP832974-A2.
 XX
 PD 01-APR-1998.
 XX
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX
 DR WPI: 1998-195228/18.
 DR N-PSDB; AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA;

Query Match 86.7%; Score 2344.5; DB 19; Length 579;
 Best Local Similarity 78.8%; Pred. No. 1.3e-196;
 Matches 435; Conservative 32; Mismatches 34; Indels 51; Gaps 1;

OY 1 MKPTSLIMASAGALALAAPAAQVTPYTDLLANPAGEWISYQONQENRHSPLTQIT 60
 DB 1 mkptslimasagalalAAPAAQVTPYTDLLANPAGEWISYQONQENRHSPLTQIT 60
 OY 61 TENVGQLOLVWARGKQKQVOTPLIHGQVWYLANPGVIOAIDAKTGDLIWEHRQLPN 120
 DB 61 tenvgqlqlvwarqmgpkqvotplihdgvywylanpgdivaiddaktdglwehrrqlpn 120
 OY 121 IATLNSFG-----I 129
 DB 121 iatlnsfgeptrgmalylgtvnyfsvwdnhlvaldtatgvtfdvdrqggedmvsnsqpi 180
 OY 130 VANGYVAGSTCOYSPFCEVSGHDSATGEBELNRNFIFRAGEGDEMGNGNYEARMTG 189
 DB 130 vangyivagstcoyspfcevsghdsatgelnrnfifragegdegmngnyearmtg 240
 OY 181 vanyvlgvsgtscqspgctvsgndsatgeelwnyflpraseegdetvngdyearmtg 240
 DB 181 vanyvlgvsgtscqspgctvsgndsatgeelwnyflpraseegdetvngdyearmtg 240
 OY 190 AMGOITTPVNTLVHGSTAVGASFTQGTGEGTLYGTNTRFARVPDTEIWRHQLP 249
 DB 241 awgiltvdpvntlvhgstavgasetqgtgtygtlgntrfavrptdteiwrrhqlp 300

QY 250 RDNMDQECTFEMAVTNDVOPSTMEGLQSIINPAATGERRVLTGVCKTGMQFDAET 309
 DB 301 rdnmdqectfemavtndvopstmeqlslnpaatgerrrvltgvpcktgimqfdaet 360
 QY 310 GEFLARNTYQNMIESIDENGIYVNEADILKELDYEDVCPFTFLGGRMPSAALNPDS 369
 DB 361 geelwardtsyenllesidengivrvdeskylteltdpvcpllllgyridwpsaalnpd 420
 QY 370 GIYFIPLNNVCYDMAVDOFTSMQVYNTSWTKLPCKDMIGRIDADISTGRTIMSV 429
 DB 421 giyfiplnncmdieavddefslavynstakmapkclvgrldidstgrtlwtiae 480
 QY 430 RAAANYSPIVSTGSGVLENGTDRYFRALSOETGETLMOTRLATVYASQALSYEVDMOY 489
 DB 481 reaanyspvlstagsvlfngtdrfralsgetlwtqlratvasqavsyeldgvy 540
 QY 490 VAIAGGVSYGS 501
 DB 541 laigsgltlys 552

RESULT 3

AAW37876 AAW37876 standard; Protein: 579 AA.

AAW37876;

10-AUG-1998 (first entry)

Alcohol and/or aldehyde dehydrogenase B amino acid sequence,

Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulonic acid; L-ascorbic; inhibition.

Glucanobacter oxydans.

Location/Qualifiers

Key 1..23

Peptide /note="signal peptide"

Protein 24..579

/note="mature protein"

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOPE) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB: AAV29054.

Recombinant glucanobacter oxydans alcohol and/or aldehyde

dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1; Pages 44-46; 59pp; English.

This is the amino acid sequence for the Glucanobacter oxydans

alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes

or recombinant organisms can be used to convert suitable substrates

to aldehydes, ketones or carboxylic acids, especially to convert

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be

converted to L-ascorbic acid by standard procedures. The derivatives

of AABH enzymes have desired substrate specificity, higher affinity

to a substrate, lower affinity to an inhibitory compound, higher

CC stability against temperature and/or pH and higher catalytic speed.

XX Sequence 579 AA;

Query Match 85.4%; Score 2310; DB 19; Length 579;

Best Local Similarity 76.5%; Pred. No. 1,5e-195;

Matches 426; Conservative 38; Mismatches 41; Indels 52; Gaps 1;

1 MKPTSLMASAGALALLAPAFQVTPVDELAPNPPAGEWISYQONCENTRHSPLQIT 60

1 mnpfllrltsaavllltpaataqvpidellampagewinygrnqenyrhsplqit 60

61 TENYGOLOLVARQMPQKVQVTPPLIHGVMTLANGVDTQADKGTGLIWEHRQLPN 120

61 adnvgqldlvaigmeagavgvpmldhgywylanpgdvlgaldqtdllvehrrqlpa 120

121 IATLNSFG----- 128

121 valtnaagdkrkqvalygltslyfsswdnhlialdmetgqvfdvergsgedqltsntg 180

129 IVANGVIVAGSTCOYSPGCFVSGHDSATGHELMRNKFTIRAGEBDEGMNDYEARMT 188

181 ivangvivasgicqspygcfisgdsatgeelwrnhfipqpeggetvgnfearmt 240

189 GAWGQITYPVTNLVHGSTANGPASEFQRCPTLYGNTREAVRPTGEIWRKHQTL 248

241 gvwgqitlypvtnlvfygstangpasecqrtpgltlygntlrfaivrptdgeiwrhql 300

249 PRDNMDQECTFEMAVTNDVOPSTMEGLQSIINPAATGERRVLTGVCKTGMQFDAE 308

301 prdnmdqectfemavtndvopstmeqlslnpaatgerrrvltgvpcktgimqfdaa 360

309 TGEFLARNTYQNMIESIDENGIYVNEADILKELDYEDVCPFTFLGGRMPSAALNP 368

361 tgeflardntynqmiesidengivrvdeskylteltdpvcpllllgyridwpsaalnp 420

369 SGIFYFIPLNNVCYDMAVDOFTSMQVYNTSWTKLPCKDMIGRIDADISTGRTIMSV 428

421 tgyfiplnncmdieavddefslavynstakmapkclvgrldidstgrtlwtiae 480

429 ERAANYSPIVSTGSGVLENGTDRYFRALSOETGETLMOTRLATVYASQALSYEVDMO 488

481 eraanyspvlstagsvlfngtdrfralsgetlwtqlratvasqavsyeldgvy 540

489 YVAIAGGVSYGSGLNS 505

541 ylaigsgltlygtqlna 557

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOPE) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB: AAV29054.

Recombinant glucanobacter oxydans alcohol and/or aldehyde

dehydrogenase A'' enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulonic acid; L-ascorbic; inhibition.

Glucanobacter oxydans.

Location/Qualifiers

Key 1..23

Peptide /note="signal peptide"

Protein 24..578

/note="mature protein"

EP832974-A2.

XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 PF 19-SEP-1996; 96EP-0115001.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjo M, Tomiyama N.
 PI WPI; 1998-195228/18.
 DR N-PSDB; AAV29053.
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 CC Claim 1; Pages 41-43; 59pp; English.
 PS This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC Sequence 578 AA;

Query Match 83.2%; Score 2251; DB 19; Length 578;
 Best Local Similarity 75.7%; Pred. No. 2,5e-190;
 Matches 420; Conservative 37; Mismatches 46; Indels 52; Gaps 2;
 QY 1 MKPSILMASAGALALIAFAFAVPTVDELAMPAGEWISYCONENRHSPLQIT 60
 DB 1 mktllqssaallvlgltlalaq-latdemlampagewinygqenyrhplcqt 59
 QY 61 TENVGOLQIWMARGOMPKQVOTPIHOGVMYLANPCVIOALDAKTGDLWEHRROLPN 120
 DB 60 adhvqglqvwatgmeagkivtrpvhqgmymyianpavdigaalcaagdlwehrqlpn 119
 QY 121 IATNSFG-----I 129
 DB 120 iatlnsfgeptgmalylgtlnvfyvswdnhlvaldstgqvfvtdrgqgtdmvsnsaprl 179
 QY 130 VANGVIVAGSTCOYSPFCFVSGHDSANGSELMNRYFIPRAGEBDEFTWGNDEYEAHMTG 189
 DB 180 vanylvagstcgyqpfscfsgndsaageelwtrnflpragegdetwgndeyeanmtg 239
 QY 190 AMGOITPYNVLVHGYSTANGPASEMORGTGGTLYGTTRAVRPPDGEIVMKRQTLP 249
 DB 240 wqgqlytbpvvglyhlytsaavpaeeiqrgtvgmsygntrfavregeelvwthqqlp 299
 QY 250 RDNDDECTEFEMVNTNDVOPSTEMEGLSINPNATGERRVLTGPKCTGTMMOFDAET 309
 DB 300 rdndgectefemvnyvndvpsaemeglnahnpdaatgerivvtgpcnglmwgfdaet 359
 QY 310 GEFLMADTYNOMIESIDENGIVTNEDAILKEKDEVEDVCPPLGGRDMPALNDS 369
 DB 360 geflwaedtyqnllesvdpdglvhnedlvleleavayelcpllggrdmpsalnldt 419
 QY 370 GIYFIPLNVCYDMAVDOETSMOYNTSNTKLPKPKDMIGRIDALIDISTGRTLSVE 429
 DB 420 giylfiplnacsgmtavdegsldvnyvsldklspssemgrididistgltlsae 479
 QY 430 RAANASPVLTSGGVLFWNGTDRFRLASOETGTLTQTLAVASGOALSTEVGOMQY 489
 DB 480 ryasnvpavltsggvlfwngtdrfyrlasgetgtlwtqtrlatvasgaalsteyldgvy 539

QY 490 VAIAGGVSVSGSLN 504
 DB 540 vaiigrvgtsygsnln 554

RESULT 5

ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 DF 01-MAR-2001 (first entry)

DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulononic acid.
 XX Gluconobacter oxydans.

PN W0200065066-A1.
 PD 02-NOV-2000.

PF 23-APR-1999; 99WO-1B00736.
 PR 23-APR-1999; 99WO-1B00736.

PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.

PI Choi E, Rhee S, Lee E;
 XX WPI; 2000-687351/67.
 DR N-PSDB; AAC83153.

PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PS production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 PS Claim 1; Fig 8; 96pp; English.

CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell.
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH sequences (fragments of the SDH subunit
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.

XX Sequence 754 AA;

Query Match 17.8%; Score 480.5; DB 21; Length 754;
 Best Local Similarity 26.2%; Pred. No. 1.8e-33;
 Matches 157; Conservative 80; Mismatches 234; Indels 129; Gaps 17;

QY 14 LALLAFAFAVPTV-----TDELLAMPAGEWISYCONENRHSPLQITTEVNGQ 66
 DB 17 lgcacaalafatspvalaedtgatlnedqhpqdwmsygrtyseqryspldqfckdnasn 76
 QY 67 LQIWMARGOMPKQVOTPIHOGVMYLANPCDVIOALDAKTGDLWEHRROLPN--NIA-- 122
 DB 77 lklawhydlatnrggeqtdplivdyvmyatnwskmkdaatgtklwsvdpkypgnladr 136
 QY 123 -----TLNSEFIVANGVIYAGS-----TCQYS 144

[illegible]

RESULT	10
ABG24430	
ID	ABG24430 standard; Protein; 443 AA.
XX	
AC	ABG24430;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #24421.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	W0200175067-A2.
XX	
PD	11-OCT-2001.
XX	

PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YF;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS88617.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID No 54789; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers or a food supplement. (II) and its binding partners are useful in medical disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAG00010-ABC30377 represent novel human diagnostic amino acid sequences. AAG00010-ABC30377 represent novel human Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 443 AA;

Query Match 5.7%; Score 155.5; DB 22; Length 443;
 Best Local Similarity 31.5%; Pred. No. 4.7e-05;
 Matches 46; Conservative 21; Mismatches 50; Indels 29; Gaps 7;

QY 6 LTMAS-----AGALALAAPAFQVPTVDLILANPPAGEWISYGNONENYRHSPLTQ 58
 DB 168 ltwagfndpgeingltlsadatpaea-ispvadt-----dwpaygrnqegqitfslpkq 218
 QY 59 ITTENVGOLQVMA-----RCMOPGKV--QVTPLIHDGVMYLANPEDVIOAIDAKTGD 109
 DB 219 inadhvmlkkaawftrgdvkgpndpeltnevtprkvgtdlylctahqrlfaiaaasgk 278
 QY 110 LIMEHRQLNINATLNSF-GIYANGV 134
 DB 279 ekwhyd---pelktneisqhvtrcgv 301

RESULT 11
 AAM09614
 ID AAM09614 standard; Protein; 824 AA.
 AC AAM09614;
 XX
 DT 20-MAY-1997 (first entry)
 XX
 DE Pathogenic Staphylococcus aureus ORF1 gene product.
 XX
 KW Pathogenicity; vaccine; immunoassay; antibody.
 XX

OS Staphylococcus aureus strain NCTC 8325-4 (ISP8).
 XX
 PN W09641878-A1.
 XX
 PD 27-DEC-1996.
 XX
 PF 13-JUN-1996; 96WO-AU00353.
 XX
 PR 13-JUN-1995; 95AU-0003507.
 XX
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Adham El-Adhami W, Matthei KI, Stewart PR.
 XX
 DR WPI: 1997-065465/06.
 DR N-PSDB: AAT47757.

Nucleic acids specific to pathogenic Staphylococcus aureus strains for diagnosis and immunogenic protein prodn.
 Claim 30; Page 62-65; 88bp; English.

An 824-amino acid protein (AAM09614) and 192-amino acid protein (AAM09615) were identified as products of 2 open reading frames in a nucleic acid (AAT47757) isolated from pathogenic Staphylococcus aureus strain ISP8. The 2 gene products are highly hydrophilic and have high surface probability and antigenic index measurements, suggesting that they are surface expressed. The polypeptides, their fragments, and antibodies raised against them can be used in vaccines and as immunoassay reagents for detecting pathogenic S. aureus strains.

Sequence 824 AA;

Query Match 5.0%; Score 134.5; DB 18; Length 824;
 Best Local Similarity 19.3%; Pred. No. 0.0087;
 Matches 88; Conservative 58; Mismatches 131; Indels 179; Gaps 20;

QY 19 APAFAQV-----PVTDELLANPPAGEWISYGNONENYRHSPLTQITTEVNGOLQV 70
 DB 427 apgtelkvregqkgekliprptlknpltgelskgeskeelkdpinel----- 476
 QY 71 WARGMOPGVVPLIHD-----GWYLANPGDVIOAIDAKTGDLIWEHR 115
 DB 477 -----eygpetilprhndefdklptgekeevpkygklnp-----etgdiv----- 518
 QY 116 ROLNINATLNSFQIVANGVIVAGSTCOYSPFCFVSGHDSATGEBLMRNFIIRAGEED 175
 DB 519 -rppvdsavkygypvkgdsivekeelpfkkerkf--npdlapgtck-----vteggkge 569
 QY 176 ETWGNDEYKRWMTG-----AMGQITDPVTNIVHGSTAVGPA-----SE 215
 DB 570 ktlitptlknpltgelskgeskeelkdpinelteygpeliprhnfdelkplptgeke 629
 QY 216 TORCPGTGLTGNTRFAVRPT-----GEIWRHQTLPRDNMOOECTFEAMATN 265
 DB 630 evpkykpy--lknpekgdvrvpvdsvckypvkgdsivekeelp--fekerkf-----n 679
 QY 266 VDVQVSTE-----MEGLQSI----- 280
 DB 680 pdlapgtelkvregqkgeklitpklknpltgelskgeskeelkdpinelteygpel 739
 QY 281 -----NPNMATGGERVVLGVPCRTGTMMOFDAETGELFARDTNVONNIESIDNGI 332
 DB 740 tprhnfdelkplptgeke--evpkygkik--npeitgdiv-----rppvdsavkyg- 785
 QY 333 VTVNEDAILKELDVEYDVCPFTFLGGRDMSALNDP 368
 DB 786 -pvkgdsivekeelp-----kerkfnd 809

RESULT 12

CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 910 AA;

Query Match 4.9%; Score 133.5; DB 22; Length 910;
 Best Local Similarity 25.1%; Pred. No. 0.012;
 Matches 97; Conservative 46; Mismatches 151; Indels 93; Gaps 23;

QY 26 TPVTDELLANPPAGEMISYQONQENYRHSPLQITTT---ENV-----GQLQVWARGMQ- 76
 DB 231 tptt-----knplgkvggeptektkpvdeltgysgeelkpgkhdefpnabpksge 286
 QY 77 --GK-----VOYPLIHGVMYLANPGDYIAIDAKTGDLIWEHRRQL-PNIA TL 124
 DB 287 dvpgkpgvknptdgvvtpvddvdkypvddgpi-----tsteelpfdkkrdefpnlap- 341
 QY 125 NSFQIVANGVIVAGSTCOYSPFCGVSCHDSATGEEIMRNYPFRAGEEGDEDTWGNDEYA 184
 DB 342 gtekvvgkg--epgkklitp-----tknplgkvg-----gegept----- 377
 QY 185 RMTGAMGOITVDPTNLVHYGSTAVGPA--SETQRGTPGTLGTMTRFAVR-PDTGEI 241
 DB 378 -----ekvtkqpvdelvhyggeelkpgkhdefdnabpksgevpkpgvknptdgv 430
 QY 242 VMRHQTLPNDMDQECTFEEMAVTN---VDVOPSTEMEGL-----OSINPNAATG-ERRVL 292
 DB 431 v-----tpvdd-----vltkypvddgplsteeelpfdkkrdefpnlapgtekvgv 476
 QY 293 TGVF-CGT-GTMMQFDETFEFLMARDTNYQNMIESIDENGITVVEDAILKELDERYDV 350
 DB 477 kgpgtklittptknplgkvggeptektkpvdv-----lvhyggeelkpgkhdefdp 534
 QY 351 CPTFLGGRDMPWSAAL-ALANPDSGIYFIP 375
 DB 535 napkysgedvpgkpgvknptdgvvtp 561

RESULT 14

AAG82914
 ID AAG82914 standard; Protein: 696 AA.

XX AC AAG82914;
 XX DT 03-SEP-2001 (first entry)
 XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.
 XX KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX KW vaccination; endocarditis.
 XX OS Staphylococcus epidermidis.
 XX PN WO200134809-A2.
 XX PD 17-MAY-2001.
 XX PF 09-NOV-2000; 2000WO-US30782.
 XX PR 09-NOV-1999; 99US-0164258.
 XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Kimmery WJ;
 XX DR WPI: 2001-316495/33.
 XX DR N-PSDB: AAH53764.
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -

XX
 PS Claim 18; Page 763-764; 2188bp; English.

XX AAH52304 to AAH5970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG81520, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH55098 represent oligonucleotide sequences from the present invention.
 CC N. B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 696 AA;

Query Match 4.9%; Score 133; DB 22; Length 696;
 Best Local Similarity 23.9%; Pred. No. 0.0092;
 Matches 106; Conservative 57; Mismatches 152; Indels 128; Gaps 27;

QY 26 TPVTDELLANPPAGEMISYQONQENYRHSPLQITTT---ENV-----GQLQVWARGMQ- 76
 DB 103 tptt-----knplgkvggeptektkpvdeltgysgeelkpgkhdefdnabpksge 158
 QY 77 --GK-----VOYPLIHGVMYLANPGDYIAIDAKTGDLIWEHRRQL-PNIA TL 124
 DB 159 dvpgkpgvknptdgvvtpvddvdkypvddgpi-----tsteelpfdkkrdefpnlap- 213
 QY 125 NSFQIVANGVIVAGSTCOYSPFCGVSCHDSATGEEIMRNYPFRAGEEGDEDTWGNDEYA 184
 DB 214 gtekvvgkg--epgkklitp-----tknplgkvg-----gegept----- 249
 QY 185 RMTGAMGOITVDPTNLVHYGSTAVGPA--SETQRGTPGTLGTMTRFAVR-PDTGEI 241
 DB 250 -----ekltkqpvdelvhyggeelkpgkhdefdnabpksgevpkpgvknptdgv 302
 QY 242 VMRHQTLPNDMDQECTFEEMAVTN---VDVOPSTEMEGL-----OSINPNAATG-ERRVL 292
 DB 303 v-----tpvdd-----vltkypvddgplsteeelpfdkkrdefpnlapgtekvgv 348
 QY 293 TGVF-CGT-GTMMQFDETFEFLMARDTNYQNMIESIDENGITVVEDAILKELDERYDV 350
 DB 349 kgpgtklittptknplgkvggeptektkpvdv-----lvhyggeelkpgkhdefdp 405
 QY 351 CPTFLGGRDMPWSAAL-ALANPDSGIYFIP 375
 DB 406 -----pnabpksgevpkpgvknptdgvvtpvddvdkypkvgvnpstteel 455
 QY 395 -----VYNTSNVTKLPFGKDMI 411
 DB 456 pfdkkrvfnpd-----lkpggeev 474

RESULT 15

AAU34349
 ID AAU34349 standard; Protein: 1016 AA.

XX AC AAU34349;
 XX DT 14-FEB-2002 (first entry)
 XX DE Staphylococcus aureus cellular proliferation protein #625.

XX Antisense: prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

XX N-PSDB; AAS52208.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids

XX Example 3; Seq ID No 5845; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1016 AA;

XX Query Match 4.9%; Score 133; DB 22; Length 1016;

XX Best Local Similarity 21.2%; Pred. No. 0.016; Indels 184; Gaps 27;

XX Matches 113; Conservative 59; Mismatches 176; Indels 184; Gaps 27;

XX 19 APFAQVY-----PVDELANPPAGEMWISYONONRHSPLTQITENYGOLOLV 70

XX 528 apgtekvregqkgekltpklkpllgvliskgeskeeltkdpinelt----- 577

XX 71 WARGQPKQVYPLIHD-----GVWYLANPGDVYQALDAKTDGLIWEHR 115

XX 578 -----eypetlcpghndefdklptgekeevpgkpkpkn-----etgdyv----- 619

XX 116 ROLPNIATLNSFGIVANGVYAGSTCOYSPFGVSGHDSATGEELRWYFLPRAGEEGD 175

XX 620 --rppdvstkygpyvkgdivekeelpfekektf--npdlapptek-----vtrgdkye 670

XX 176 ETWGNDEYKARMTGA-----WQIITYDPVTNLVHGSTAVGPA-----SE 215

Db 671 ktltpklkpllgvliskgepkeeltkdpineltgypetltpghndefdklptgeke 730
 QY 216 TORRPGSTLYGTNRRAVRPDT-----GELVNRHOTLPDRDNDODCTFEAMVTN 265
 Db 731 ewpgkpy--lknpetgdvvrpvdsvtkygvkgsdivekeelp--fkertkf-----n 780
 QY 266 VDVQSTEMEGLSINPMAATGERRVLTGVPCCKGTMMQFDATGTEFLMARTNYONNME 325
 Db 781 pdlappte-----kvtrgqkgektitf--ptlknpl-----tgeil-----skye 819
 QY 326 SIDENGIVTNEADAILKELDVEYDVCPFLGGRDWPASAL-----NPSDG 370
 Db 820 sice-----ltdkpl-nel--teygpeltpghndefdklptgekeevpgkpkpknpetg 872
 QY 371 IYPLPANNVCYDMMADVDEFTSMQVNTSNVTKLPCK-DMIGRIDALIDSTGRITMSVE 429
 Db 873 dvvrp-----pvd-----svtkygvkgsdivekeelpfeke 905
 QY 430 RAAANYSPLVSTGGVLFNGGIDRFRALSOETGELMOTRLATVASGOAIS 481
 Db 906 ----kfnplap-----gtekvtrg-qdkgektitltpklkpllgvlis 944

Search completed: May 24, 2002, 10:19:20
 Job time: 383 sec

Fri May 24 11:27:28 2002

walick-934-128.pcp.rag

Fri May 24 11:27:29 2002

walick-934-128.pep.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:08 ; Search time 39.53 Seconds
(Without alignments)
312.040 Million cell updates/sec

Title: WALICK-934-128.PEP
Perfect score: 1 MKPTSLMASAGLALALAP.....GMOYVALAGGVSYSGSLNS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	480.5	17.8	754 4 US-09-296-284-4	Sequence 4, Appl1
2	479	17.7	738 1 US-07-985-458-3	Sequence 3, Appl1
3	471.5	17.4	720 4 US-09-296-284-25	Sequence 25, Appl1
4	319	11.8	739 2 US-09-136-251-2	Sequence 2, Appl1
5	106	3.9	966 2 US-08-868-786-2	Sequence 2, Appl1
6	105	3.9	1012 1 US-08-219-262B-10	Sequence 10, Appl1
7	105	3.9	1012 2 US-09-031-655-10	Sequence 10, Appl1
8	103.5	3.8	974 2 US-08-868-786-6	Sequence 11, Appl1
9	101.5	3.8	1042 3 US-08-928-361B-11	Sequence 16, Appl1
10	101.5	3.8	1837 3 US-08-928-361B-5	Sequence 16, Appl1
11	99.5	3.7	1381 4 US-09-540-245A-16	Sequence 4, Appl1
12	99.5	3.7	3567 2 US-07-642-734C-4	Sequence 4, Appl1
13	99.5	3.7	3567 3 US-08-439-009A-4	Sequence 4, Appl1
14	97	3.6	543 4 US-08-856-841-22	Sequence 22, Appl1
15	97	3.6	1721 4 US-08-700-651-5	Sequence 5, Appl1
16	96.5	3.6	606 2 US-08-883-534-3	Sequence 3, Appl1
17	96.5	3.6	606 3 US-09-204-764-3	Sequence 3, Appl1
18	95.5	3.5	548 2 PCT-US93-10541-2	Sequence 2, Appl1
19	95.5	3.5	500 2 US-08-987-519-2	Sequence 2, Appl1
20	95	3.5	816 1 US-08-229-444B-2	Sequence 4, Appl1
21	95	3.5	816 2 US-08-541-780-4	Sequence 4, Appl1
22	95	3.5	816 3 US-08-216-276A-19	Sequence 19, Appl1
23	95	3.5	1012 1 US-08-714-402-2	Sequence 2, Appl1
24	95	3.5	1112 2 US-09-356-818A-2	Sequence 2, Appl1
25	94.5	3.5	512 4 US-08-462-484-10	Sequence 10, Appl1
26	94.5	3.5	527 1 US-08-441-147-10	Sequence 10, Appl1
27	94.5	3.5	527 1 US-08-441-147-10	Sequence 10, Appl1

28	94.5	3.5	527 5 PCT-US95-07536-10	Sequence 10, Appl1
29	94	3.5	635 4 US-08-931-608A-5	Sequence 5, Appl1
30	94	3.5	824 4 US-09-626-589-3	Sequence 3, Appl1
31	93.5	3.5	548 1 US-08-247-902A-2	Sequence 2, Appl1
32	93	3.4	380 4 US-08-971-782-4	Sequence 4, Appl1
33	93	3.4	380 3 US-09-309-026-4	Sequence 4, Appl1
34	93	3.4	459 3 US-08-971-782-2	Sequence 2, Appl1
35	93	3.4	459 4 US-09-309-026-2	Sequence 2, Appl1
36	92.5	3.4	659 4 US-09-626-589-2	Sequence 1, Appl1
37	92.5	3.4	717 4 US-09-626-589-1	Sequence 1, Appl1
38	92.5	3.4	5215 4 US-09-105-537-2	Sequence 2, Appl1
39	91.5	3.4	532 3 US-08-737-336-6	Sequence 12, Appl1
40	91.5	3.4	954 4 US-09-251-645-12	Sequence 2, Appl1
41	91.5	3.4	1612 1 US-08-169-927-2	Sequence 6, Appl1
42	91	3.4	500 2 US-08-987-519-1	Sequence 12, Appl1
43	91	3.4	551 2 US-08-793-229-32	Sequence 32, Appl1
44	91	3.4	551 3 US-09-285-957-32	Sequence 32, Appl1
45	91	3.4	682 3 US-08-481-435-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucosyltransferase Suboxylans Sorbitol Dehydrogenase, Genes
; FILE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: 1533.08/0000
; CURRENT APPLICATION NUMBER: US/09/296.284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Glucosyltransferase suboxylans
; US-09-296-284-4

QY	14	LALLAAPFAQVTV	17.8%; Score 480.5; DB 4; Length 754;
QY	17	LGCMAALAFCTSPVALAEDTGTATINADQHPDMSYGRYSSEORVSPDIOITKDNASN	Best Local Similarity 26.2%; Pred. No. 1.5e-37; Indels 129; Gaps 17;
QY	67	LQVWAROMQGVQV-TPLIHGQWYLANPQVLOAIDAKGDLIWEHRRLP-NTA	Matches 157; Conservative 80; Mismatches 234;
QY	77	LILAHYLDLDRQEGPLIVGVYATYTNWSKRALDAATKILMSYDQKVPENIADR	
QY	123	TLNFGIVANCVIYAGS	
QY	137	GCDDYVNRGAAYNKNVFFGTFGRDLALDAKGLIWSVYTPKRAOLGHRSTYVNGA	
QY	145	P-FVSGHDSATGEELMRNYEFPRAEGED	
QY	197	PLRAKKVILIGGAEFGARGVATAYDAETGKMDRFTVPNDKRPDGAASDYLMKSA	
QY	176	ETWGDYERKMTG-AWGQITDPTNTLVHGSTAVPASSETORGTGPGGTLTYTNR	
QY	257	YPTWGGKAGAKQGGGGGYWDSLIIDPYTDLVYLGVGSGSWNKFKPSECKGNLFLGSI	
QY	232	FAVPPDTGELIWHQHTLPDNNNDQECTEDMAVNTYVDPSTWEMGLQISINPAATGE-RR	
QY	317	VAIMPDTGKXVWHFQFTPMDDMDYTSVQDIALMDMPV	

```

QY 291 VLTGVCCTGTMOQDAETGEFLMARDTNYNMIESIDE-NCITVNDALIKELDEYD 349
Db 360 VLVHAP--KNGFYIIDAKTGKFTISGRPYTEWMANGLDVPGRPVNDALMTLNGKPMY 418
QY 350 VCPTELGSDMPASALNPDGSIYFIPLNVCY-----DMVAVOQETSMDY 336
Db 419 GIPDGLGHNFAAASPOKRLVYIPAOQVPPVYDPOKGFKAHDSMGLDMNKIGL 478
QY 397 NTSN----VTKLPKGMIGRIDDAIDISTGRILMSVERAANYSFVLSTGGVLFNGGTD 452
Db 479 DDNDPOHKADKQOFLDKLGVAMPDPOKQAFVYDHKGFWMGGLATAGVLFQGLAN 538
QY 453 RYFALSQETGETIMOTRLATYASGOAISYEVDGMQYVAI-----AG-----GGVSYGSG 502
Db 539 GEHHAIDATTCGDLFFPPQASAIIPVPTYANGKQYVAEVGMGSIYFPLGQVARTSG 598

```

RESULT 2

US-07-985-458-3
Sequence 3, Application US/07985458

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takemura, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing the Same And Transformed Acetic Acid
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fishaut, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
OPERATING SYSTEM: IBM PC compatible (NEC PC-9801 ES)
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about

```

IDENTIFICATION METHOD: 72,000
ORIGINAL SOURCE:
ORGANISM: Acetobacter altoacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Takemura, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horinouchi, Suenaru and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

```

Query Match 17.7% Score 479; DB 1; Length 738;
Best Local Similarity 25.9%; Pred. No. 2e-37;
Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

```

QY 10 SAGLLALAAPFAQVTPYTD-----LLANPAGEWISYGOQNEVYRHSPLQIT 61
Db 17 TACTICALISGATMASADGQATGEALIHADHDHGNMTYGRYSQKRSPLDQNR 76
QY 62 ENVGQLQVWARGQKGVQV--TPLHDGVWYLANPQVDYQAIDATGCLINERROL 119
Db 77 SNVGNLKLAWYDLDTNRGQGTPLVDGVYATYTWMSMKKAVDAATGRLMSYDPRV 136
QY 120 NIA-----TLNS-----FGI----- 129
Db 137 NIADGCCDVTYRGAAYNGKRYFGTFDGLILDAKTGLVWSVNTIPPEALGKORSY 136
QY 130 -----VANGVIVAGSTCOYSPFGC--FVSGHDSATGELMRNFIPRAGEBD 175
Db 197 TYDGAPRIAKGRVYIGN--GGSEFGARGFVSAPFDETGKVDREFTYVNPKNEDASDS 254
QY 176 -----ETWGNDEYKRMKTA-----WQITVDPVTNLVHYGSTAVGASPTOR 218
Db 255 VLMNKAYQWS-----PTGAWTRQGGGYWDSIYDPAVLVYLGNGSPWYKTR 307
QY 219 GTPGGLYGTNTRFAVRPDTGETVWRHQTLPDNDQDECFEMAVTVNDVQSPTEMG 278
Db 308 SEGKGNLFLGSIYALKPRTGEYVNHFOETIPMDQDFTSDQITMTLDTLP----- 357
QY 279 SINPNATGERRYLVGPKCTGTMOQDAETGEFLMARDTNYNMIESID-ENGITVNE 337
Db 358 -----NGETRVIVYHARKNGEYFIIDAKTGFTISGRPYTEWMANGLDVPG 410
QY 338 DALKEIDVEYVCPTELGSDMPASALNPDGSIYFIPLNVCY-----DMVAVO 396
Db 411 DALYTLTGKREYIGIPDGLGHNFAAASPOKRLVYIPAOQVPPVYDPOKGFKAH 478
QY 397 NTSN----VTKLPKGMIGRIDDAIDISTGRILMSVERAANYSFVLSTGGVLF 452
Db 479 DDNDPOHKADKQOFLDKLGVAMPDPOKQAFVYDHKGFWMGGLATAGVLFQGLAN 538
QY 446 LFGSGTDRYFALSQETGETIMOTRLATYASGOAISYEVDGMQYVAI----- 495
Db 531 LFOGLANGEFHAYDATNGSGLFFHADSGIIPVPTYANGKQYVAEVGMGSIYFPL 598

```

RESULT 3
US-09-296-284-25
Sequence 25, Application US/09296284A

```

; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: 1533 0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 720
; TYPE: PRF
; ORGANISM: Glucobacter suboxydans
; US-09-296-284-25

```

```

Query Match 17.4%; Score 471.5; DB 4; Length 720;
Best Local Similarity 26.0%; Pred. No. 1e-36; Indels 125; Gaps 17;
Matches 151; Conservative 81; Mismatches 224;

```

```

26 TTTTDELLANPAGWISYGNQENYRHSPLTQITTEVGOLOLVMARQMPKQYV-TP 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   5 TATTN---ADHPGDMSTGRTYSEGRYSPLDQITKDNMSNLKAMHYDLDTRNQBEGTP 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   85 LHDGVMTLANPGDVIOAIDAKGDLWEHRQOLP-NIA-----TLNFGVANGVIA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   62 LIYDGVMTATWMSKMLDAATGKLMSYDPKPGNADRGCCDVNNGAAYWNGKYVF 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   138 GS-----TCQYSP-----FCG 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   122 GTFDGRLIALDAKGLVMSVYTPKEAQLGHRSTYVDGAPRIAKGVIIINGCAEFRA 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   149 --FVSGHDSATGELMRNFTIRAGEBD-----ETWGNDEYARMWTG---A 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   182 RGFYATADAEETGKMDRFEFTVPRPNKPDGAASDVLSKAPYTWGSGAMKQGGGTV 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   191 WQGTTPRVNVLVHGSTAVGASSETORCTPGTLYGNTFRAPVPTGELVYRHOITLP 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   242 WDSLITPDVLDVLYGVNGSPWNYKFRSEKGNMFLGSIVAIAPDGLGHNFAAATSPQ 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   251 DNMDOCTEFEMAVTNDVDPSTMEGLQSIINNAATGE-RVYLGVPCKTGTMMQFPAET 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   302 DQMDVTSVOQIMALDMPV-----NGEMRHVLAAP-KNGFFYIIDAKT 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   310 GEFLLARCTNYQNIESIDE-NGIVTVNEDAILKELDVEYDVCPTFLGGRMPAALNPD 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   344 GFFISGKPYTYENNANGIDPVTGPRNYPDALWTLNGKFWYGIPIGDLGHNFAAATSPQ 403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   369 SGIVFPIANNY-----DMAVDOEFTSMVYNTSN---VTKLPFGKDMI 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   404 TKLVYIPAOVFFVYDPQKGFKAHDSWNLGIDMANKIGLLDNDQHKADAAQKLLK 463
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   412 GRIDALDISTGTLMSVERAANAYSPVLSGGVLENGSDTRFKRLSDETEGTLMOQL 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   464 GMLVANDPQKQAAFTVYDHKGPMMNGGLATAGVLFQGLANGEFHAYDATGKDLTPRA 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   472 ATVASGAISTEVDGMQYVAI---AG-----GGVSYSGS 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   524 OSALIAPVYTYANGKOVAYEVGNGIYPPFLGCVARTSG 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-09-136-251-2
; Sequence 2, Application US/09136251A
; Patent No. 6127156
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: MIYAZAKI, Taro
; APPLICANT: OJIMA, Setsuko
; APPLICANT: SHINJOH, Masako

```

```

; APPLICANT: TOMIYAMA, No. 61271561bun1
; TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
; FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
; CURRENT APPLICATION NUMBER: US/09/136,251A
; CURRENT FILING DATE: 1998-08-19
; EARLIER APPLICATION NUMBER: EP 97114432.4
; EARLIER FILING DATE: 1997-08-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 739
; TYPE: PRF
; ORGANISM: Glucobacter suboxydans
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(24)
; US-09-136-251-2

```

```

Query Match 11.8%; Score 319; DB 3; Length 739;
Best Local Similarity 21.8%; Pred. No. 5.5e-22;
Matches 149; Conservative 73; Mismatches 197; Indels 266; Gaps 32;

```

```

23 AQVTPVDELLANPAGWISYGNQENYRHSPLTQITTEVGOLOL--VMARGM--OPG 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   86 SVPAPAPQDSANPARQDWAYGRDDHQRTYSPLEITPENASKLKAFTVHTGSTRPG 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   79 KY-----QVTP-L-IHDGVMTLANPGDVIOAIDAKGDLWEHR-----ROLPNATLNSF 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   146 QYNKMAETTPKIDGDLGYCSAMNDITK-LDPATGKQIWRNVDVKYHSIPYPAACKGV 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   128 GIVANGVIVASTCOYSPFGCFVSGH-----DSATGEELMRNFTIRAGE----- 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   205 TYFTSSVYVEGPCH-----NRLIGTLDMRLIAVDAETD-----FCPNFGHGQVNL 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   173 -----EGDET-----GNDYEARMTGAM--- 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   255 QGLGESVPGFVSMTAPPVYINGVYVNHVYLDGRRAPSGVIRGYDASGFFVAMOVN 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   192 ---GQITPDVTLNVLHGSTAVGASSETORCTPGTLYGNTNR-----FAVRPD 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   315 NSGRSPAT-RVYITAVERRIPGLPDRRORG-----GSRLLPDRNSAADYYSALRSD 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   238 -----TGEIVMRHOITLPDNDDOCTEFEMAVTNDVQVSTEMEGLQSIINP 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   367 AENKVSAAVAIDVKTGSPRWVFOYAHKDYMDYDLSGATL-----MDMPG---P 413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   283 NAATGERRVLTGVPCKTGTMMQFPAETGEFLMARDTNYQNIESIDEN-----GIV--- 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   414 DGQIVPALIM--PTKRGQTFVLDRTGK-----PILVEERPAAPSPCVIPGD 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   334 -----TVNEDAI---LKELD-----VEY 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   460 RSPQPMVSGMPALRVDPDLKETDMKMSPLDOLFCRIKFRANVYGEFTPPSVDPKPIEX 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   349 DVCPTFLGGRMPAALNPDSCIYFIPLN-NCYCDMAVDOEFTSM----- 393
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   520 ---PQYNGSDMGMSYDPOGILIANMNTIPWYDOLVTRKADSGLMPIDDPNFKRG 576
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   394 -----DYVNTSNVTKLPCKDMIGRIDALIDISTG-RILMS--- 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   577 GGAEGANGAMDTPYGIYVTPWDOY-TGMCMNRP---YGMITALDMKHGOKVLMOHL 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   428 -----VERAANYSPVLSGGVLENG-CTDHFYFALSDETEGTLMOQL 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   632 GTARANGPWGLPTGLPWEIGTPNNGSVYTGGLIFGATDNOITAIIDENHGKVVMSAV 691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   471 LATVASGAISTEVDGMQYVAIAG 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   692 LPGGQANPMTYEANGHOYVAIAG 716
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5

```

US-08-868-786-2
 ; Sequence 2, Application US/08868786
 ; Patent No. 5998701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kawchuk, Lawrence M.
 ; APPLICANT: Armstrong, John
 ; APPLICANT: Lynch, Dermot
 ; APPLICANT: Knowles, Richard
 ; TITLE OF INVENTION: Potatoes Having Improved Quality
 ; TITLE OF INVENTION: Characteristics and Methods for Their Production
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5730 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80803
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/868,786
 ; FILING DATE: 04-JUN-1997
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/036,946
 ; FILING DATE: 10-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sullivan, Sally A
 ; REGISTRATION NUMBER: 32,064
 ; REFERENCE/DOCKET NUMBER: 8-97
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303) 499-8080
 ; TELEFAX: (303) 499-8089
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 966 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-868-786-2

Query Match 3.9%; Score 106; DB 2; Length 966;
 Best Local Similarity 19.2%; Pred. No. 0.31;
 Matches 87; Conservative 56; Mismatches 154; Indels 156; Gaps 19;
 QY 25 VPPVDELIANPPAGWISGQNGENYRHSPLQITTEVNGOLV-----MARGM 75
 DB 613 VTPRWIRFCNPPLSATITWTGTEDW-----VTKTERLAELOKFNADNEDLQNEWREAR 666
 QY 76 QGVVQVYPLIHGCVMTLANPGDVIQIDAKTGDIWEHRRLPMLATNSFGIVANGVI 135
 DB 667 RSKIKVVSFLKERTGTSVYP-DAMPDIQVKR-IHEKRLQ-----LNFIV----- 712
 QY 136 VAGSTCYSPFGCVSGCHSATGGEELMRNYFIPRAGEGDETWGNDYEA---RWMTGAM 191
 DB 713 -----YRKKM-----KEMTAERKTN-FVPRVCIJFGKAFATYVQAKRIYKFTITDG 759
 QY 192 GQITTD-----PVTNLIYHVSAAVGPASEQR---GTPG--- 222
 DB 760 ATTNHDPPEIGDELKVFVDPYVNSVAELLIPASDLSEHISTAGMEASGSNNKFAWNGCI 819
 QY 223 --GTLYGTWTRAVRDPDGEIVWHRQLPRLPNMDQECTFEMAVTVNDVQSTEMEGLQSI 280
 DB 820 QIGTLDGANVE--IREEGE----- 837
 QY 281 NNNAATGERVLTGVPCKTGTMMQFPAETGEFLMARDTNYQNMNISIDENGIVYVNEAI 340
 DB 838 -----ENFFLFGAQAHEIAGLKRERADGKY--PDEFEEVKEFVNSGAFGSYNDDL 888

QY 341 LKELD-----VEYVCPFLGGRDMPAALNDSGIYFIPLMWCYDMKMAVDQFTSM 393
 DB 889 IGSLEGNEGRGRADY-----FLVGRKDFPS-----YIEQGEVDEAYIDQKRWITM 933
 QY 394 DYVNTSNVTKLPKPGKDM-----IGRIDAIDIS 420
 DB 934 SILNTAGSYKFSDDRTIHEYAKDLMWNEAVEIJA 966

RESULT 6
 US-08-219-262B-10
 ; Sequence 10 Application US/08219262B
 ; Patent No. 5788970
 ; GENERAL INFORMATION:
 ; APPLICANT: VAKHARIA, VIKRAM
 ; APPLICANT: SNYDER, DAVID B
 ; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 ; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
 ; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 ; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/219,262B
 ; FILING DATE: 29-MAR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 2747-047-27
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1012 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Infectious bursal disease virus
 ; STRAIN: OH
 ; US-08-219-262B-10

Query Match 3.9%; Score 105; DB 1; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.42;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;
 QY 199 VTNLYHVSSTAV-----GPAS-----ETORGTPGGLTGTWTRF 232
 DB 1 MTNLMDDHQIIVPFRISLMLPTGPASIDPDILEKHTLRSETSTYVNLVSGTGLIVFF 60
 QY 233 AVRPDTGEIVWHRQTLPRD---NMDOECTFEMAVT-----NVVOPSTE 273
 DB 61 PGFP--GSVGAHTTLOSNGSYQFDQ-----MLTPAQLMPSYVNCRLVRSLSLTVRSSTL 113
 QY 274 MEGLOSINP--NATGERRVLTGVPCKTGTMMQFPAETGEFLMARDTNYQNMIESI----- 327

Fri May 24 11:27:29 2002

walick-934-128.pep.rat

Page 5

Db 114 PGGVALNGTINAVT-----FQGSISEL---TDYSYGLMSATNIN 152
QY 328 DENGIVTNEDEALILELDEYDVCPTFLGRDPSKALNPDGSIYFIPLNNVC-----380
Db 153 DKIGNVLVEGVTYLSLPTSYDLSTYRLGD-PIPAAGLD-----KLMATCDSRDR 204
QY 381 -YDMAVADQ-EF-----TSMDEVNTSNVTKLPKPKDM-GRIDAIDISGRTLM- 426
Db 205 VYVTAADEYQFSSOLIPSGVKTILFTANIDALSLSVGGELIFSQTTHSLEDVYTIYF 264
QY 427 -----SVERAANYSPVLSSTGGVLFN--GGDRYFRALSOETGETLMOTRLATYAS 476
Db 265 IGFDTETVYKAVATDFGLTTGTNNLVPENLGPTSETIQPITSMKLEVYTKRGGT--A 322
QY 477 GOAISYEVDGKQYVAIAGG 495
Db 323 GDPISMTVSGTLAVTIYVG 341

RESULT 7
US-09-031-655-10
; Sequence 10, Application US/09031655
; Patent No. 6017759
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SANDER, DAVID B
; APPLICANT: MENGEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,655
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/219,262
; APPLICATION NUMBER: 29-MAR-1994
; FILING DATE: 29-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: OH
; US-09-031-655-10

Query Match 3.9%; Score 105; DB 3; Length 1012;
Best Local Similarity 22.2%; Pred. No. 0.42;

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;
QY 199 VTNLVHKGSTAV-----GPAS-----ETORGTPGCTLYGTNTRF 232
Db 1 MTNLMQHTQIYPIFRSLMPTGASIPDDTLEKHTLRSEISTYLVLDGSGSLYVF 60
QY 233 AVRPDGEIYWRHQTLPD-----NMDQCTFEMAVT-----WVDQPESTE 273
Db 61 PGFP--GSVGAHYTLQSGSYQFPO-----MLLTAONLPVSYNYCRLYSRSLTYRASSYL 113
QY 274 MEGIOSIMP--NMAIGERYLGLVCKTGCTMMQFDEAFGEFLMARDTNQNMTEST- 327
Db 114 PGGVALNGTINAVT-----FQGSISEL---TDYSYGLMSATNIN 152
QY 328 DENGIVTNEDEALILELDEYDVCPTFLGRDPSKALNPDGSIYFIPLNNVC-----380
Db 153 DKIGNVLVEGVTYLSLPTSYDLSTYRLGD-PIPAAGLD-----KLMATCDSRDR 204
QY 381 -YDMAVADQ-EF-----TSMDEVNTSNVTKLPKPKDM-GRIDAIDISGRTLM- 426
Db 205 VYVTAADEYQFSSOLIPSGVKTILFTANIDALSLSVGGELIFSQTTHSLEDVYTIYF 264
QY 427 -----SVERAANYSPVLSSTGGVLFN--GGDRYFRALSOETGETLMOTRLATYAS 476
Db 265 IGFDTETVYKAVATDFGLTTGTNNLVPENLGPTSETIQPITSMKLEVYTKRGGT--A 322
QY 477 GOAISYEVDGKQYVAIAGG 495
Db 323 GDPISMTVSGTLAVTIYVG 341

RESULT 8
US-08-868-786-6
; Sequence 6, Application US/08868786
; Patent No. 5968701
; GENERAL INFORMATION:
; APPLICANT: Kawchuk, Lawrence M.
; APPLICANT: Armstrong, John
; APPLICANT: Lynch, Dermot
; TITLE OF INVENTION: Knowles, Richard
; TITLE OF INVENTION: Potatoes Having Improved Quality
; TITLE OF INVENTION: Characteristics and Methods for Their Production
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: 5730 Manhattan Circle, Suite 201
; STREET: Boulder
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80803
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,786
; FILING DATE: 04-JUN-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,946
; FILING DATE: 10-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Sally A
; REGISTRATION NUMBER: 32,064
; REFERENCE/DOCKET NUMBER: 8-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 449-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 974 amino acids
; TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-868-786-6

Query Match
Best Local Similarity 3.8%; Score 103.5; DB 2; Length 974;
Matches 89; Conservative 65; Mismatches 166; Indels 161; Gaps 19;

14 LALAAPAAQVTPYDELLANPPAGE---WISTQONEN-----YRHSPLQIT 60
Db 579 LCVSGHVAAGVAIHSIIVDEVFNEFYKIMPEKQNKNTGVTFRMLSCNELSEII 638
QY 61 TENVC-----OQOLVARGKOPKQVQVPTLHDSVMTLANGDV 99
Db 639 TKWTGSDMLVNTKELALEKRPADNELOSEWRKAKGNKKKIYSLIKETGYVSP-DA 697
QY 100 IQADAKTGLIMEHRQPLNATLNSFGIVANGVYAGSTCOYSPGCFVSGHDSATGE 159
Db 698 MEDVQIR---IHEYKROL-----LNFQIV-----YRKKN-----KEMSP 731
QY 160 ELARNYFIPRAGEGDETWGNDYEA---RMKGAMQOITYDPV-----199
Db 732 EERKEKVPVPCIFGSKAPATYVQAKRIKFTIDYGETVNHDEIGDLKVFVVDIYVS 791
QY 200 -----TNLVHGSTAVGAPSETOR---GTPG---GLYGTNTRFAVRPDGETIYWR 244
Db 792 VAEVLIPGSELQSHISTAGMEASGTSNMKFSMNGCILLITDGAWE---IREEVEGE---845
QY 245 HQTLPRDNMDOCTFEMMAYTVNDVQSTEMEGLOINPAANGERRVLTVGPKCTGTMMQ 304
Db 846 -----DNFFLEGAQAEHIGLRKREAEKGFVDDPFEEVKAFITGV---FGTY--891
QY 305 FPAETGEFLMARDNTYONMIESIDENGITYVNEDAILKELDEYDVOPTLGGRDMPASA 364
Db 892 -----NYEELMGSLEGN-----EGGRADYFLVGKDPD--920
QY 365 LNPDSGIYFPLNNVCTDMAVDOEFTSMDYNTSNVTKLPKPGDM-----IGRIDAID 418
Db 921 -----TIECODKVDAYVRQKKTKMTSLNLTANGSFKFSORTIHOYARDIMRIEPE 972
QY 419 I 419
Db 973 L 973

RESULT 9

US-08-928-361B-11

Sequence 11, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn

TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

TITLE OF INVENTION: SPECIES INFECTIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETERS, VERNY, JONES & BIKSA

STREET: 385 Sherman Avenue, Suite 6

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,361B

FILING DATE: 12-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, HANA
REGISTRATION NUMBER: 30,518
TELEPHONE/DOCKET NUMBER: 480,76-1(HV)
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1042 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-11

Query Match
Best Local Similarity 3.8%; Score 101.5; DB 3; Length 1042;
Matches 138; Conservative 71; Mismatches 222; Indels 229; Gaps 36;

25 VTPVDELLANPPAGEMISTYQONDE-----NYRHSPLQITTEWYQOLVARGKOP 77
Db 301 IDPTTG-LFPNPPGHLINPTNNMTDSFAGAYKTAVANGIKTIDVYGLPVDITGLPK 389
QY 78 GKQVQVPLHDSVMTLANPGDVIAIDAKTGLIMH-----RROLPNATLN-----125
Db 360 DPVSDIP-----FNSYTGEL---VDPSTGKPINNYAGIVSGKRGILPDEDENGLDE 409
QY 126 ---SEGIYANGVI-----VAGST---COYSPGCFVSGHDSATGEELARNYFIPRAG 171
Db 410 PSTKLPIQNGNQLVNPETNSTYVSGTSGTKRP-GIPVNGGVVPEE-----AK 459
QY 172 EGGDETMNGNIEARWMTGAMGOITTPVNTLVHGSTA---VGPASEYQKSTG-----222
Db 460 DQAK--GKD---GLIVPPTNSINRDPVTNQTNTGNTGNIINP---ETGKVIPLSGSLPS 512
QY 223 -----GLYGTNTRFAVRPDGETIYWRQTLPRDN---WDQECTFEMAYTN 265
Db 513 YPSFNTPOQTDITGKPDVDTVGLPYDPSGELIDPATKLPILPGSVAGDELTEVANTT 572
QY 266 -----VDVQ---PSTMEGLQS-----INP-----NAAG-----ERRV 291
Db 573 DEVTGLPIDLETGRLPRDVSGLPOLPNGTLVDPSSNKKPIPGSHSGFINGISGEQSHKDP 632
QY 292 LTGVPCCKTGMKQPD-----AETGEFLMARDNTYONMI---ESIDENGIYVNEDAIL 341
Db 633 STGKFLDNTGLPDEDGSLINPETGDKLQSGHSGTTPMPVPGKPOENCGIMTPPO-IL 691
QY 342 KELDY---EYDQCP-----TFE---GGRDMPSEAL-----365
Db 692 EALKLFTLSNEVINSRPPSDAVPDRPTNTWNNKISGTFQVQCKKITIGSASVYIHAL 751
QY 366 -----NDSGIIYFPLNNVCTDMAVDO-----EFTSMDYNTSNVTKLP 406
Db 752 GTPYQIDPTTGLSDPSGTGLPIPGFNVLVDPQGEIKSVYVSLYKKNITTEAY 811
QY 407 GKDMIGRIDAIDISTGRULMSVERAANYSPVYSTG-----GVLFNGGIDRYFRA---457
Db 812 GL-----PVPKGTGPIIDI-----STLPKKNELIDPISGKYFSGAGIFISKAG 859
QY 458 ---LSOETGETIMQT-----RLATVASGQAISEYDGMQYVAIAG-----GGV 497
Db 860 SQSKSDSGNPIIDPSTNMPYDPKTKLIDPESGIALINDSVSGV-FATVGTAPKPKGV 918

RESULT 10

US-08-928-361B-5

Sequence 5, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 TITLE OF INVENTION: SPECIES INFECTIONS
 NUMBER OF INVENTIONS: 30
 CORRESPONDENCE ADDRESS:
 ADDRESS: PETERS, CARYN, JONES & BIKSA
 STREET: 385 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: VERNY, HANA
 REGISTRATION NUMBER: 30,518
 REFERENCE/DOCKET NUMBER: 480.76-1(HV)
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1837 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-5

Query Match 3.8%; Score 101.5; DB 3; Length 1837;
 Best Local Similarity 20.9%; Pred. No. 2.4; Indels 229; Gaps 36;
 Matches 138; Conservative 71; Mismatches 222

25 VPEVTDELLANPPAGEMISYQNOE-----NYRHSPLQITTEVNGOLQVWARGMOP 77
 1096 IDPTTG-LFPNPTGLINPTNNMTDSSFGAYKYKAVSNGIKTDNYGGLPVEITGLPK 1154
 78 GKQVPELIDHGVMIYAIADAKTGLIMEH-----RROLNATATLN----- 125
 1155 DPVSDIP-----VDSGTGEL---VDPSTGKPINNYAGIYSGKRGLPREDENGLMFD 1204
 126 ---SFGIVANGVI-----VAGST---CQYSPGCEVSGHDSATGELMRNTFIRAG 171
 1205 PSTYLPIDGNNOJVNPEITNSVSGSTSKPKP-GIPVNGGAGVYPDE-----AK 1254
 172 EEGDEITGNDYEARMGTGANGQITTYDPTVNLVHGSTA---VGPASERGRGPG----- 222
 1255 DQADK---GKD---GLIYPTNSINKDPVINTQYSNTGNTINP---ETGKVIPGSLPGSLN 1307
 223 -----GLYGTNTRFAVAPDPGELIYMRHQTLPDRN---MDQECTEEMVYN 265
 1308 YPSENPQOQDEITGKPVYDTGTPDPSTGELIDPATKLPISVAGDEILTEVLNITT 1367
 266 -----VDVQ---PSTEMEGLOS-----INP-----NATG---ERRV 291
 1368 DEVTGLPIDLETGLPRDPVSGLPDLPNGTLVDSNKKPIPGSHSGFINGTSEQSHKDP 1427
 292 LFGVCKTGTYMOPD-----AETGEFLMADTYNOMNI---ESIDENGIVTVNEDAIL 341

1428 STGKPLDPNTGLPEFDEDSGLINPETGDKLGSGHSTFMPVGRQGENGIMTPEO-IL 1486
 342 KELDY-----EYVCP-----TFI---GGDWPSSAAL----- 365
 1487 EALNKLPSTSEVVISRPPSSDAPDRPTWNNKISGOTFOVDGKKTIPGSAASVHTAL 1546
 366 -----NPDSGIYFIPLNVCYDMMAVDQ-----EFTSMOYVNTSVTKLP 406
 1547 GTPQOTDPTTGLSDSPSTGLPFIQGFNVLDVPOTEQIGSVYVSLYKREKNTVEANAY 1606
 407 GKMDIRIDAIDISTGRTLMSVERAANYSPLYSTG-----GGVLFNGIDRYEFA--- 457
 1607 GL-----PYDPKGTGPIDPT-----SLPFAKNGBELDIPISGKFFSGSINAF 1654
 458 -----LSOETGETLMOT-----RLATVASGQAIISYVDQMTVAIAG-----GGV 497
 1655 SOSKSSDSGNPDPSTNMPYDPRTGKIDPESGIAIDNSVSGV-FATVGTAPARKGGV 1713

RESULT 11
 US-09-540-245A-16
 Sequence 16, Application US/09540245A
 Patent No. 6270984

GENERAL INFORMATION:
 APPLICANT: Goodman, Corey
 APPLICANT: Kid, Thomas
 APPLICANT: Brose, Katja
 APPLICANT: Tessier-Lavigne, Marc
 TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 FILE REFERENCE: 898-031-3
 CURRENT APPLICATION NUMBER: US/09/540,245A
 CURRENT FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 60/065,544
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: 60/081,057
 PRIOR FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 1381
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-540-245A-16

Query Match 3.7%; Score 99.5; DB 4; Length 1381;
 Best Local Similarity 20.6%; Pred. No. 2.4; Indels 125; Gaps 20;
 Matches 90; Conservative 51; Mismatches 170

97 GDVIO-----AIDAKTGLIME-----HROLNATATLNSFGIYANGVIYAGS 139
 626 GDVIELNASVVDSTSMKLTWQIINGKYVEGFYVAROLPNPIVNNPAPVSNTPPLGS 685
 140 TCQYSPGCEVSGHDSATGELMRNT-FIRAGEEDETMGN-----DYEARMT-- 188
 686 TS-----TSASASASASALISTKPNIAAKRGELINOSGAGAPPPLNRYMLTIL 737
 169 ---GANGQITTYDPTVNLVH-----GSTAVGASETOGRGPGTGLGTMR 231
 738 NCGGASCT---ITGIVQYTYLFEFFIYFYSYEGKRSNRJARLTLEDVPSRAPYGEA- 793
 232 FAVRPDTEIYWRHQTLPDRMDOCTEFEMMYTNVDOPSTEMEGLOSINPNAATGERRV 291
 794 LILNSSAVFLMKAPEL-KDRHGYLVNTHVYKID-----TANHSRI 836
 292 LFGVCKTGTYMOPDAETGEFLMADTN---YONMIESIDENGL-----VYNEDAILK 342
 837 LINV-----TIDAASPTLVLANLTEGVMTYVGAAGNNAVCYCPATRLDPLTK 888
 343 ELDEVYDCPTLGRD-----WPSAALNPDGSIYFIPLNVCY---DMAAVDOE 389
 889 RLD-----PRINORDHVNDVLTQPFITLLGAILAVIMLSGAMVYVRKRMMAKQGA 941

QY 390 FTSMDVYNSNTKLPCKDMIGRIDALDISTGRTLM-----SYERA---AANYSPVL 439
 Db 942 LNTMGNHTSDVKMPSLSARNGCYWLDSTGGMWRPSPGGDSLDMOKDHIADYAPVC 1001
 QY 440 ----STGGVLENGCT 451
 Db 1002 GAPSPGAGTSSGGS 1017

RESULT 12

US-07-642-734C-4
 ; Sequence 4, Application US/07642734C
 ; Patent No. 5824513
 ; GENERAL INFORMATION:
 ; APPLICANT: Katz, L
 ; APPLICANT: Donadio, S
 ; APPLICANT: Mcalpine, J B
 ; TITLE OF INVENTION: Recombinant DNA Method for Producing
 ; Erythromycin Analogs
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Edward H. Gorman
 ; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 ; STREET: Park Rd
 ; CITY: Abbott Park
 ; STATE: IL
 ; COUNTRY: US
 ; ZIP: 60064-3500
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/642,734C
 ; FILING DATE: 17-JAN-91
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dancigers, Andreas M
 ; REGISTRATION NUMBER: 32652
 ; REFERENCE/DOCKET NUMBER: 4952.US.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 708-937-9396
 ; TELEFAX: 708-938-2623
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3567 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-642-734C-4

Query Match 3.7%; Score 99.5; DB 2; Length 3567;
 Best Local Similarity 22.9%; Pred. No. 11;
 Matches 125; Conservative 48; Mismatches 221; Indels 153; Gaps 26;

QY 21 AFAQVTPVVDLILANPPAGEW-----ISYQNOENYRHSPLTQITTEVNGQLDVLW 71
 Db 2911 AFAPIAVTDBRLARVPDG-WSDADAANVPIAY-----TTHAYLHDLACLRAGQSVLIH 2964
 QY 72 ARGMQPGKQVYPLIHDSGYMYLANPGDYIOAIDAKTGLIWEHRRQI-----PNATLNSF 127
 Db 2965 AAGGCGMAAVALARRAGAEVILATAG-----PAKHGTL-----RALGLDEHILASGRF 3013
 QY 128 G-----IVAN--GVIVAGSTQYSPFGCFYS-CHDSATGEELMRNFIP 168
 Db 3014 GFARKFRERTGRGVYVILNSITGLLDESADLLAEDGVFEWKGKTDLRDAGDGRGRYAP 3073
 QY 169 -RAGEGSDTWNNDY-ERAMMTGANGQITTYDPTNLY-----HYGSTAV- 210
 Db 3074 FDLGEGDRLRGLREIVVYGLGA-GEILDRLPVSAMELGSAPALQMSGRHVGKLVLT 3132

QY 211 GPASRTQRT---PGTLYGINTFRAVPDRTGELVWRHQT-PRDMNDQECTFEAMVTVN 266
 Db 3133 QPAPVDDPDTVLITGCT--GTLGRILARHLVTEHGVHLLVSRGADAPGSELAELE 3190
 QY 267 DVQPTSTEMEGLQSIQNPNAF-----GERVLTGVCCKTGMQWDATGELTARDNTION 322
 Db 3191 DLGASAEIACDPTADRDALSALLDGLRPLTGVVHAAGV----- 3230
 QY 323 MIESIDENGIVT-VNEDAILKELDEVYDVCPTFLGGRDPSALNDPSGIYFIPINVCY 381
 Db 3231 -----ADGLVTSIDEPAVEQVLRKVDAA-----NWLHET/ANTGLSFVLFSSNA 3276
 QY 382 DMAAVDQFTSMQDYNTSNV-----TKLPCK-----DMGR 413
 Db 3277 SVLA-----GPGQVYAAANESLNALALRTRRLPAKALGWLMAQASEMTSGLDRIAR 3332
 QY 414 IDAIDISTGRILMSVERAANYSPVLSTGGVLENGCTDRYFRALSOETGETL-WQTRLA 472
 Db 3333 -----TGVAALPTERALLFDSALRRGGEVYFPLISIRSLARRAEVPEVLRGMVRAK 3385
 QY 473 TVASGOA 479
 Db 3386 LRAAGQA 3392

RESULT 13

US-08-439-009A-4
 ; Sequence 4, Application US/08439009A
 ; Patent No. 6004787
 ; GENERAL INFORMATION:
 ; APPLICANT: Donadio, S
 ; APPLICANT: Katz, L
 ; APPLICANT: Mcalpine, J B
 ; TITLE OF INVENTION: Method of Directing Biosynthesis of
 ; Specific Polypeptides
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Steven F. Weinstein
 ; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 ; STREET: Park Rd
 ; CITY: Abbott Park
 ; STATE: IL
 ; COUNTRY: US
 ; ZIP: 60064-3500
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/439,009A
 ; FILING DATE: 11-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Casuto, Dianne
 ; REGISTRATION NUMBER: 40,943
 ; REFERENCE/DOCKET NUMBER: 4952.US.D1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 847-938-3137
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3567 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-439-009A-4

Query Match 3.7%; Score 99.5; DB 3; Length 3567;
 Best Local Similarity 22.9%; Pred. No. 11;
 Matches 125; Conservative 48; Mismatches 221; Indels 153; Gaps 26;

QY 21 AFAQVTPVVDLILANPPAGEW-----ISYQNOENYRHSPLTQITTEVNGQLDVLW 71

Fri May 24 11:27:29 2002

walick-934-128.pap.ra1

Page 9

DB 2911 AFAPLVTHRLARLAVPDG-MSDADAAVPLAY-----TTAHYALHDLGLAGLAGSVLTH 2964
QY 72 ARGMOGKQVTPPLHGVMTLANPGDVIAIDAKTGLDIMEHROL-----PNIATLNSF 127
DB 2965 AAGVGMAVALLARRAGAEVATAG-----PAKHGTL-----RALGDDEHLSSTRET 3013
QY 128 -----IVAN-----GVYAGSTCQYSPFGCVS--GHSATGEELMRYVFP 168
DB 3014 GFARKFRERTGRCGVVNLSTLGLDLESADLLAEDGVFEAKGTDLRAGDFRGRTAP 3073
QY 169 -RAGEGDETFWMDY--EARMGTGAMGOLTYDPVNLV-----HYGSTAV- 210
DB 3074 FDLGAGDRLGELLREVVGLGA--GEDRLRVSAMELGSPALLOHNSRGRHVKLVLT 3132
QY 211 GPASETORG---PGTLVGTNTREAVRPTGELIYWRHQL--PRDNWDOECTFEEMVTNV 266
DB 3133 OPAPVDPGTVLITGT--GTLGRLLARHLVTEHGVRLHLLYSRGGADAPGSDLELRAEIE 3190
QY 267 DVQSTEMEGLOSTNPAAAT-----GERRVLTGVCKTGMOPDAETGEFLMADNTYON 322
DB 3191 DIGASAEIAACDTADROLASLDGLRPLRGVHAAGVL----- 3230
QY 323 MIESIDENGIVT-VNEDAILKELDEYDVCPTFLGSDMPSALNPDGSIYPIPLNNVCY 381
DB 3231 -----ADGLVTSIDEPAVEQVLRAKVDA-----WMLHLLFANTGLSFFVLFSSAA 3276
QY 382 DMAVNDQEFSDVYNTSNV-----TKLPKG-----DMIGR 413
DB 3277 SVLA---GGGCVYAAANESLNALALRRRGLPARALGWLMAQASEMTSLGDRIRAR 3332
QY 414 IDAIDISTGRTMSVERAANSPVLSGGVYENGDTFRYALSOETGETL-WOTRLA 472
DB 3333 -----TGYAALPTERLALALFDSLRKGEVYFPLSIRNSALRRAEVFEVLNGMVRAR 3385
QY 473 TVASGOA 479
DB 3386 LRAAGOA 3392
RESULT 14
US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US/08/214,658
APPLICATION NUMBER: 16-MARCH-1994
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22
Query Match 3.68; Score 97; DB 4; Length 543;
Best Local Similarity 21.0%; Pred. No. 0.9; 168; Indels 146; Gaps 23;
Matches 96; Conservative 48; Mismatches 168; Indels 146; Gaps 23;
QY 41 WISYGNQENYRHSPLT-QITTEVNGOLQVWARGMOPKQVTPPLIHGVNYLA----- 94
DB 96 WRYLDNFNALNLFSPLEPHQHLIENYGI-----APDALTY-----ISRIANKVY 141
QY 95 -NPGVYIADAKTGD--IWEHROL-----NIATL 124
DB 142 DKGGGVQVDTSTGRCLMVDHEKRYVLAGGQDTLAPELPIWYFPOVAYLVTVGV 201
QY 125 NSRGIVANGVYAGSCQYSPFGCVSGHS-----ATGEELMRYVFPFRAGEG----- 174
DB 202 NTGIGSDSKKLSSESAF-----VLEHSSPOLLGTGTSMSKXKPPRPPEPLESCSQ 256
QY 175 -----DEWGNDEYEAARMGTGAMGOLTYD-----PVNULVYVYSTAVGPA 213
DB 257 HFYEMYNPLYSRLGVDPDLGDPKRSLSLHEDHALOPQNFMPGFLVNSV--STKEGDS 313
QY 214 SETORGTP-GGTLVGT--NTRFAVRPTGELIYWRHQLTRDNWDOECTFEEMVTNVV-- 266
DB 314 SNTGAKALTLGISTGTSQNTSIRLP--GPV-----SQYTHHMDT-----DKVYTGINAIIS 362
QY 269 -----QSTEMEGLOSTIN-----PNAAGIRRVLTGVVCKTGT 301

Db 363 HGOITTYGNAEDKEXQGVGRFPEKQOLKOLGMLNHTYFPKNGTQOYTDQIERPLAVGS 422
 QY 302 MWOPDATERGE-IMAR-----DINYQNMIESIDENGIVAVNEDALLKELDE-IDVCEPTEL 355
 Db 423 VMNRRAHYSQMSKIPMLDSEFKTOFALGCMGLHPPOITLKOYAVGIMTYMTERK 482
 QY 356 GGRWPSAALNPDGSIYFIP-----LNNVCYDMMAYD 387
 Db 483 LGPRKATGRWNPQGVY--PRAACHLPYVLYDPLAD 518

RESULT 15
 US-08-700-651-5
 ; Sequence 5, Application US/08700651B
 ; Patent No. 6015882
 ; GENERAL INFORMATION:
 ; APPLICANT: PETERSEN, CAROLYN
 ; APPLICANT: LEECH, JAMES
 ; APPLICANT: NELSON, RICHARD, C.
 ; APPLICANT: GUT, JIRI
 ; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
 ; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
 ; FILE REFERENCE: 480.19-4(HV)
 ; CURRENT APPLICATION NUMBER: US/08/700,651B
 ; EARLIER FILING DATE: 1997-08-14
 ; EARLIER APPLICATION NUMBER: 08/415,751
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1721
 ; TYPE: PRT
 ; ORGANISM: Cryptosporidium parvum
 ; US-08-700-651-5

Query Match 3.6%; Score 97; DB 3; Length 1721;
 Best Local Similarity 20.5%; Pct. No. 5.9;
 Matches 136; Conservative 78; Mismatches 215; Indels 234; Gaps 38;

QY 25 VTPVDELINPAGENISYQNOE-----NYRHSPLQTTENNGQLQVLVARGMOP 77
 Db 979 IDPTTG-LPFPNPTGHLINPTNNTMDSSFAAYKAYVANGIKTDNV-----YGLPV 1029
 QY 78 GKVOVPLIHGVMYLANPGVIAQDAKTDGLIWEHRRLPNATLNSFGIVA----- 131
 Db 1030 GEITGLP-----KDPGSDI--FPNSTGTGLVDPSTGKPKNNSTA--GIYSGKPLP 1076
 QY 132 -----NGVI-----VAGSTQYSP-FGCFVSGHDSAT-----GELMRNFTIP--RA 170
 Db 1077 PIEDENGNI.FDPSYMLPIDGNNOLVNPETNSTVSGSTGTTKPKPIPVNGGVVPEDEA 1136
 QY 171 GEGDETWGNDYEARWMTGANGOLITDVPVNLVHIGSTA--VGPASETORGTPG----- 222
 Db 1137 KQOAKK--GKD-----GLIYVPTNSINKDPVNTQYSMITGNIINP--ETGKVIPLSLPGSL 1189
 QY 223 -----GLIYGINTRFAVPRDPTGEIWMRHOTLPRDN--WDQECTFEMAVT 264
 Db 1190 NYPSPNTQOQDETITGKPVDTVGLPYDPTSTGEIIPATKLPPIPSVAGDELITVELNIT 1249
 QY 265 N-----VDVQ-----PSTEMEGLS-----INP-----NAATG-----ERR 290
 Db 1250 TDEVYGLPIDLETGLPRDPSGLPOLPNCTLVDPSPNKKPIPSHSGFINCTSGEQSHED 1309
 QY 291 VLTGVPCKTGT--MMOPD-----AETGEFLMARDTNY-----QNM 323
 Db 1310 PSTGKPLDPTGILHPDEDSGLINPETGDKLOGSHSGTFMVPVGRKQGENGIMTPEQI 1369
 QY 324 IESIDENGIVTVE-----DAI-----LKLDEYEVDCPTFLGG----- 357
 Db 1370 LEAL--NKLPTSNENVISPRPSDAVDRPTNTWMTKISGTYOVDGKTKIILGSAASYIH 1427

QY 358 -----RMPSSAAL--NPDGSIYFIPLNNVCYDMAAVD-----EFTSMQVYNTSNYTK 403
 Db 1428 TALGTPTQDPTTGLPSPDSGTGLPIRGFNVLVDPOTEQIKGSVPYVSLYKEKNITTE 1487
 QY 404 LPFGKDMIGRIDAIISTGRTLMSVERAANYSFVLSTG-----GVLFGNGTDRYFRA 457
 Db 1488 AAVGL-----PVDPKTGPIDPI-----SYLPFAKNELLIDPISGKYFSGSIAGFLSG 1535
 QY 458 -----LSQETGELMOT-----RLATVASSGALISYEVGMQYVAIAG----- 494
 Db 1536 KAGSOSKSSDESQNPIDPSTNMPYDPKGGKLDIPESGIALDINSVGV--FATVPGTAAPRK 1594
 QY 495 GGV 497
 Db 1595 GGV 1597

Search completed: May 24, 2002, 10:20:13
 Job time: 326 sec

Fri May 24 11:27:29 2002

walick-934-128.pep.ra1

339 AILKELDVEYDVCPTELGGRDWPSSALNFBSSA.....

Db 412 GLYTLGKFWYIGIPGLAHNFMGMAVSPETHLVLLPAHQLPFYKKNVGGFKRPHDPAW 471
 Qy 386 VDEFTSMYNTSNVTKLPPGKDMIGRIDAIDISTGRTLMSVEAANAANSPYLSSTGGV 445
 Db 472 VGLDMTKNGLPDPE-ARAYAIKIDLHGLWLLAMPVKMETVWKIDHKPMGNGVLANGDGL 530
 Qy 446 LFNKGIDRYFRALSOETGETLMOTRLATVASQAISYVDGMQVA-----IAGG 495
 Db 531 LFGGLANGEFHAYDATNGSDLYKFDAGSGITIASPWTYSVNGKQYVAVEVGMGITPISNG 590
 Qy 496 GVSXGSG 502
 Db 591 GVGRTSG 597

RESULT 2

S52317
 quinoximoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
 C:Species: Comamonas testos
 C:Date: 08-May-1995 #sequence, revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S62366; S62373; S65308; S52317
 R:Stoortvogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
 Eur. J. Biochem. 235, 690-698, 1996
 A:Title: Characterization of the gene encoding quinoximoprotein ethanol dehydrogenase
 A:Reference number: S62366; MUID:96184549
 A:Accession: S62366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STOI>
 A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
 A:Accession: S62373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STOI>
 R:de Jong, G.A.H.; Geerloff, A.; Stoortvogel, J.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
 Eur. J. Biochem. 230, 899-905, 1995
 A:Title: Quinoximoprotein ethanol dehydrogenase from Comamonas testos. Purification
 A:Reference number: S65908; MUID:95324580
 A:Accession: S65908
 A:Molecule type: protein
 A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>
 A:Experimental source: ATCC 15667
 C:Genetics: qheh
 C:Superfamily: glucose dehydrogenase (pyrroloquinoximoprotein-quinone)
 C:Keywords: oxidoreductase; quinoximoprotein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinoximoprotein ethanol dehydrogenase type 1 #status experimental <MA>

Query Match 17.8%; Score 481; DB 2; Length 708;
 Best Local Similarity 25.4%; Pred. No. 4,7e-27;
 Matches 150; Conservative 84; Mismatches 239; Indels 118; Gaps 19;

Qy 3 PSLIMASGALALLAFAFAVPTVDE-----DELANP-AGEWISYQONEN 50
 Db 12 PGRWVLLAACIG--SAAAFQTPPAQAAANVORVGDGDIRANARPTDPTIGVDAE 69
 Qy 51 YHSPULQITTEVNGOLQVYVARGMPQK-VQVTPLIHGVYMLANPGVIOAIDAKTGD 109
 Db 70 TRYSRLDQINANAKDGLAMSTYLNLESTRCVEATPVVVGIMYVANSVYHAIDTRGN 129
 Qy 110 LIWEHRRLP-----IA----- 122
 Db 130 RIMTYDIPDINSTGFKCCDVNRGVALMKKQYVYGAMDRLIALDAATGKEVHONTFE 189
 Qy 123 -----TLNSFGIANGVIAGST-CQYSPFCFVSGHDSATGELMWFNFIIP----- 168
 Db 190 GOKSLITITAPRVFKGVITIGRGAEGVRG-YTTAVDAETGKRKRKMSVSGEDSKPE 248
 Qy 169 --RAGEGDETWGNDYEAAMTGA-----WGQITDPTVNLVHYGSTAVGPASETORGTTP 221

Db 249 EDESMKRAATW--DESGKWEAGGGTMDMSMTFDEALNTMYVGNNGSPSHKVRSEK 306
 Qy 222 GGLYGTNFRFAVRPDPTGEIVRHOITLPRDWDCECFEMAMVTNVDOSTMEGLQSLN 281
 Db 307 GGDNLVIAISIALDPDTGKYKHHYOETRGDMDTSTOPMILLADIKI----- 353
 Qy 282 PNATGERVVLGVCCKTGMMQDPAETGEFLARDTNQNMIESTIDENG-LVTYNDPAI 340
 Db 354 ---AGKPRVILHAB-KNGFEFVLDRTNGKFIKNEFPVNNASGDKHGRPIGI---AA 406
 Qy 341 LKELDYEDVCPTEFLGDRDPSAALNPDGIVPIPLANNCYMAVDO-EFT----- 391
 Db 407 ARDGSKPQDAVAPGYPYGAHNNHPSFNPQGVLYLPAONPVNLMDDKWEFNQAGPKPQ 466
 Qy 392 SMQVYNTSNVTKL-PPGKDMIGRIDAIDISTRTLSVERAANAANSPYLSSTGGVLENG 450
 Db 467 SGTGWNIAKFFNAPFKSKPFGRLAMPVQAOKAMSEVHSPNNGTILTTAGNVYFOGT 526
 Qy 451 TDRFRALSOETGETLMOTRLATVASQAISYVDGMQVAIA-GGVSTG 500
 Db 527 ADGRLVYHATGEKLEAPGTGVVAPSTYMDGRQYVSAVWGCVYG 577

RESULT 3

S14270
 alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter
 C:Species: Acetobacter polyoxogenes
 A:Variety: strain NB11028
 C:Date: 21-Nov-1993 #sequence, revision 26-Jul-1996 #text_change 20-Jun-2000
 R:Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
 Biochim. Biophys. Acta 1088, 292-300, 1991
 A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
 A:Reference number: S14270; MUID:91159482
 A:Accession: S14270
 A:Molecule type: DNA
 A:Residues: 1-738 <TAM>
 A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:9216186
 A:Experimental source: strain NB11028
 C:Complex: heterodimer of 72k and 44k chains
 C:Superfamily: glucose dehydrogenase (pyrroloquinoximoprotein-quinone)
 C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MA>

Query Match 17.7%; Score 479; DB 2; Length 738;
 Best Local Similarity 25.9%; Pred. No. 7e-27;
 Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

Qy 10 SAGALALLAFAFAVPTVDE-----LANPAGWISYQONENYRHSPLQITTT 61
 Db 17 TAGTICALISGYATMASADDCGATGEALIIHADHPGMMYVGRYSDORSPLDQINR 76
 Qy 62 ENVGOLQVYVARGMPQKQV-TPLIHGVYMLANPGVIOAIDAKTGDLLWEHRRLP- 119
 Db 77 SNVGNLKLAWLIDLTNRGEGTPIVDGMYATNTWMSKRAVDAATKLLMSIDPRAVG 136
 Qy 120 NIA-----FLNS-----FGI----- 129
 Db 137 NIADGGCCDTVNRGAAYVNGKYVFGFPGDRLLAIDAKTGKLMVSVNTIPPEAELGRQSY 196
 Qy 130 -----VANGVIVAGSTQYSPFC--FVSGHDSATGELMWFNFIIPRAAGEGD----- 175
 Db 197 TVDQAPRIANGRIITN--GSEFGARGFVSFAFETGKVDWRFYVTPNKNLEPDAAS 254
 Qy 176 -----ETWGDYEAAMTGA-----WGQITDPTVNLVHYGSTAVGPASETOR 218
 Db 255 VLMNKAAYQTS-----PTGAMTRQGGGGVWDSIYVDVADLVLYGNGSPNNKYR 307
 Qy 219 GTPGGLYGTNFRFAVRPDPTGEIVRHOITLPRDWDCECFEMAMVTNVDOSTMEGLQ 278
 Db 308 SEGKGNLEFGSIYALAKPEGEYVWHFOETPRMDQMDFTSDQIMTLDLPI----- 357


```

OY 281 NENATGE-RRVLTGVPCKTGTMMQDAEGERFLARDTNQNMTESID-----EN 330
Db 339 -----GEMRHYIHP-KNGEFTYVADAKTGEISGKNVYVONNANGIDPLTRGPNVND 411
OY 331 GIYVNDALILKELDEYDYCPTEFLGGNDMPASALNPDSCGYEFTPLNNVC----- 381
Db 412 GLYTLNG-----KFWYGI-PGPIGAHNFMAAMSPKTHLYIPAIQIFGKNOVG 463
OY 382 ----DMVAVDQFTSMQVNTSNVTKLPDPRGDKMGRIDALIDISTGRTLSVERAANYS 437
Db 464 KPHADSNWVGLDWTKNGLPDPPE-ARVAYIKDLHGLMLADPVKMETWKKIDHKGPWNG 522
OY 438 VLTGGVLEFNGTDRTYFPALSOETGETLMQTRLATVASQAISEYEDGMQVYA----- 491
Db 523 ILATGDLPLGLANGEEHAYDATNGSDLYKFDAGSGIINPMTYSVNGKQYVAEVSNG 582
OY 492 ----IAGGVSYSG 502
Db 583 GIYPISMGVGVTSG 597

```

RESULT 6

```

J00706
alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
N:Alternate names: methanol dehydrogenase 62k large chain
C:Species: Methylobacterium extorquens
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
R:Accession: J00706; S07908
R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
Gene 90, 173-176, 1990
A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 mofx and mofx genes
A:Reference number: J00706; MUID:90337342
A:Accession: J00706
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-626 <AND>
A:Cross-references: GB:M31108; NID:g150017; PIDN:AAA25380.1; PID:g150018
R:Experimental source: strain AM1
R:Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989
A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
A:Reference number: S04644; MUID:89350892
A:Accession: S07908
A:Molecule type: protein
A:Residues: 28-50, 'XX', 53 <NUN>
A:Note: the source is designated as Methylobacterium extorquens AM1
C:Comment: This enzyme oxidizes methanol to formaldehyde.
C:Genetics:
A:Gene: mofx
C:Keywords: alcohol metabolism; oxidoreductase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

```

```

Query Match 16.0%; Score 434; DB 2; Length 626;
Best local similarity 24.5%; Pred. No. 1e-23;
Matches 148; Conservative 91; Mismatches 208; Indels 158; Gaps 23;

```

```

OY 10 SAAALMLA-AP-AFAQVTPYTDLLANPPAGE-WISYQNOENYRHSPLTQITENWGO 66
Db 7 SAAALMLA-AP-AFAQVTPYTDLLANPPAGE-WISYQNOENYRHSPLTQITENWGO 66
OY 67 LQIYV--ARGMQPGKVOYTPPLIHGQVW-----LANPGDV----- 99
Db 67 LQIYV--ARGMQPGKVOYTPPLIHGQVW-----LANPGDV----- 99
OY 100 -----IOAIKAGTGLIWEHRLPNIATLN 125
Db 126 RAVACCDLVNGLAYWPGDGKTPALILKQJODGNVAAALAAETGVTM----- 173
OY 126 SFGIANGVIVAGSTCOYSP-----GC-----FVSGHDSATGELMRNFEI- 167
Db 174 -----VENSIDIKVSTLTIAPIYVKRVIIIGSGAELGVAGYLTAVDYKTSQVRAIATG 229

```

```

OY 168 -----PRAGEE--GDETGWNDYEARMGTGA---WGOITYDPVTLNHYGS 207
Db 230 PDKDLLADSDENIKNPHGQGLGTGWEDG---AMKIGGCTMMGVAVADPDTNLTIFGT 286
OY 208 TAVPASETQRTPGTLYGTNTRFAVRPPDGEIYMRHOLPRDNNDECTEEMAVTND 267
Db 287 GNPAPNNTMR--PGDNKW-TMTIFGRDADTGAKEGQYATPDEWD-----YACVN 335
OY 268 VOPSTEMEGLDINPNMATGRRVLTGVPCKTGTMMQDAEGERFLARDTNQNMTESID 437
Db 336 VAMLEQD-----KDGAKRLTLHPDRNGIYTLTRTGALVANSKDDI--VAVF 385
OY 325 ESIDENGIYVNEEDAILKELD-VEYDVCPTFLGGNDMPASALNPDSCGYEFTPLNNVCYDM 383
Db 386 KSVLDKTGOVPYRDEYGTDRMDHLANDICPSAMGYNHOGHSDYDPKRELFEMGINHICMDW 445
OY 384 MAVDQFTSMQVNTSNVTKLPDPRGDKMGRIDALIDISTGRTLSVERAANYS 437
Db 446 EPFLPFRAGQFVYVATLNNYPPGKDRONEGLOIKAVNAITGDYKMEKMRFRVWGG 505
OY 438 VLTGGVLEFNGTDRTYFPALSOETGETLMQTRLATVASQAISEYEDGMQVYAIGGV 497
Db 506 TMAIAGDLYEYGTLDGYLAKRDSIDGLLMFKIPSGAIGIPMTYHKGTOYVAI----- 560
OY 498 SYSGS 502
Db 561 YGVG 565

```

RESULT 7

```

B83399
quinoprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: B83399
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-623 <STO>
A:Cross-references: GB:AE004624; GB:AE004091; NID:g9947973; PIDN:AA05370.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: exaA; PA1982

```

```

Query Match 15.4%; Score 417.5; DB 2; Length 623;
Best local similarity 23.2%; Pred. No. 1.7e-22;
Matches 139; Conservative 94; Mismatches 248; Indels 119; Gaps 19;

```

```

OY 3 PPSLWASAGALAL---LAAPFAQVTPYTDLLAN--PPAEEMISTQONENYRHSPLT 57
Db 9 PAGLLRPSLHCLAPVALGAGALAKDYTWEDIANDDKTTDVLQYGMTHAORMSPLK 68
OY 58 QTTENVGOLQLVNARM-----QPGKVOYTPPLIHGQVWLANPGDVIOAIKAGTGLIME 113
Db 69 QVNDNMYEKLTPMMSYSFGEKRGQ-ESQAIYSDGVIVYVASSRLEFALDAKAGKGLMT 127
OY 114 HRQLPN-----LATLNSGCIAN--GVIV-----AGST 140
Db 128 YNHLRPPDIRPCDDVYVNGAAYIGKVFEGTIDASVYALNNKTKVWKKRFDHAGYT 187
OY 141 COYSP-----FCGF--VSGHDSATGELMRNFEI----- 169
Db 188 MTGAPITVKGKTKVLLHSGSGDEFGVGRFLARPDPDGEIIMRPFVEHGMGLNK 247
OY 170 -----AGEBDEFTGWNDYEA-----RMWTGA--WGIITDPVTLNHYGSAVGPASET 216

```

[illegible]

Db 342 VAKVDPVAIVNEFKKLDLTGTTPVRDPEFAIRMDHKGNICPISAMGFHNGVDYSYDESRST 401
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 372 YFPIPLNNVCYDDMAAYDOEFTSMDEVNTSNVTKLDP---PKDKMIGRTIDAIDISTGRLTMS 427
| : ||::||:||||CDMEFPFLPYRAGQCFVGATLAMYGGPNGPRKKEMGOIAPFDLITTGAAKWTF 461
Db 402 LYGALNICHDMDFEPFLPYRAGQCFVGATLAMYGGPNGPRKKEMGOIAPFDLITTGAAKWTF 487
QY 428 VERAAANYSPVLSTGGCGVFNGCTDRFYRALSOEGTELMOTRIATLVASQAQSIYVDGM 521
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 462 KWEKFAMAGGTLYTKGSLWVATLDSGLKALDNKGKRLMFNKMPSSGICSPMTYSFGKR 521
QY 488 QYVALAGGVSYSGS 502
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 522 QYT-----GSMTGYG 531

RESULT 9
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli JVO107
C:Species: Escherichia coli
C:date: 30-Jun-1992 sequence revision 12-Dec-1997 #text change 16-Jun-2000
C:Accession: D64735; JVO107; A45997; A45201; I41228
C:Blahtner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <BLAN>
A:Cross-references: GB:AEO00122; GB:U00096; NID:g1786315; PIDN:MAC7325.1; PID:g17863
A:Experimental source: strain K-12, substrain WJ1655
R:Clenton-Jensen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J.Bacteriol. 172, 6308-6315, 1990
A:title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quinone A:Reference number: JVO107; MUID:91035240
A:Accession: JVO107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'RRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51523; NID:g41553; PIDN:CAA35706.1; PID:g41534
A:Experimental source: strain K12
R:Iramda, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:title: Topological analysis of quinioprotein glucose dehydrogenase in Escherichia co A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R:Pujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <NU>
A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BAAO5580.1; PID:g473791
A:Experimental source: strain K-12 substrain WJ110
R:Yamada, M.; Asakoka, S.; Salier, M.H.; Yamada, Y.
J.Bacteriol. 175, 568-571, 1993
A:title: Characterization of the gcd gene from Escherichia coli K-12 WJ110 and regula A:Reference number: I41228; MUID:93123180
A:Accession: I41228
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'RRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
A:Cross-references: GB:D12651; NID:g21655; PIDN:BAO2174.1; PID:g216556
C:Genetics:
A:Gene: gcd
A:Map position: 3 min
C:function: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone A:pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:11-37/Domain: transmembrane #status predicted <TM1>
 F:41-59/Domain: transmembrane #status predicted <TM2>
 F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93-95/Binding site: ubi quinone (Arg, Asp) #status predicted <TM5>
 F:466/Active site: Asp #status predicted

Query Match 12.6%; Score 342; DB 1; Length 796;
 Best Local Similarity 22.7%; Pred. No. 7, 5e-17;
 Matches 160; Conservative 74; Mismatches 202; Indels 270; Gaps 32;

```

QY 6 LMAS-----AGALALLAARFAQVTVYDELLANPAGEMWISGONENYRHSPLTQ 58
DB 135 LTMAGFNDPQELNGTSLADATPAEA-ISPVADQ-----DMPAYGRNDEGGRFSPKQ 185
QY 59 ITTENVGOLQLVMA-----RGMOPKV--QVPTLHDGVMYLANPGVDYQAIADAKTGD 109
DB 186 INADVHNHLEAVWFRTGDKVQPNPDEITNEVTPIKVGTLYLCTAHORLEALDAASGK 245
QY 110 LIMEHRROLPIATLNSF-----GI-----VANGYVA- 137
DB 246 EKMHYD--PELKTNESFOHTCGVSYHEAKAETASPEVADCPRIILPVNDGRILAI 302
QY 138 ----GSTCQ-----YSP-----FG 147
DB 303 NAENGKLCETFAANKGVNLQSNMPTDKPGLYEPTSPITITDKTIYVAGSVTDNFSTRETS 362
QY 148 CFVSGHDSATGEELMKNYFIPRA-----GEGDETCNDYEARMWTGANGQITDYDPT 200
DB 363 GVIKGFVNTGELLMA--FDGAKDPNAPISDEHTFTNS-----PNSWAPAAYDAKL 413
QY 201 NLVHGSTAVGPASETQRTGCTPGCLYGTNFAVRPTGETIWRHOTLPBDMDOECTPE 260
DB 414 DLVYLPKGVTTTPDIMGKNTPEQERYASSI-LALNATGKLAMSYQTVHNDLMDM---- 468
QY 261 MMTVNDVQPSSTEMEGLOSINPNAATGERVLTGVPCKTGTMMOFDAETGEFL----- 313
DB 469 ----LPAQPT-----LADITVN--GQKVPYIAPAKTGNIFFVLDNRGELVAPAPERK 515
QY 314 ---WARTNYQNMIESIDENGI-----VTNEDAILKEL--DVEYD----- 349
DB 516 VPGAAKGDVYTPQPSSELSFRPTKDLGADMGATMFDOLVCVRFHQMRKEGIFTTP 575
QY 350 ----VCPFLGGRDMPSAALNP----- 367
DB 576 SEQGLVFPNGLGMFEMGIGISVDPNREVALANPMALPFVSKLIPRGPNMEQPKDAKGT 635
QY 368 ---DSGI-----YFIPLNNVCYDMAVDOEFTSMQVNTSVTKLPBCKDMIGRIDAID 418
DB 636 GTEGSIQPOGYGVPTVTLN-----PFLS-----PGLPCCKQPMAGYISALD 676
QY 419 ISTGRLMSVERAANYS-----PV-----ISTGGVLFENGCT--DRFERAL 458
DB 677 LKTNEVVMKKRIGTPDQSMFPMPVPFPMNGMPLGISTAGNVLFIATADNLTARAY 736
QY 459 SOETGETLMOTRLATVASQA--ISYEVDGMQYVAI--AGGGSVYXS 501
DB 737 NMSGKELIMQGRLP--AGGQATPMTEYVNGKQYVVISAGHGSGFT 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A:Reference number: A95629; MUID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BB33551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC50128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.6%; Score 342; DB 2; Length 796;
 Best Local Similarity 22.7%; Pred. No. 7, 5e-17;
 Matches 160; Conservative 74; Mismatches 202; Indels 270; Gaps 32;

```

QY 6 LMAS-----AGALALLAARFAQVTVYDELLANPAGEMWISGONENYRHSPLTQ 58
DB 135 LTMAGFNDPQELNGTSLADATPAEA-ISPVADQ-----DMPAYGRNDEGGRFSPKQ 185
QY 59 ITTENVGOLQLVMA-----RGMOPKV--QVPTLHDGVMYLANPGVDYQAIADAKTGD 109
DB 186 INADVHNHLEAVWFRTGDKVQPNPDEITNEVTPIKVGTLYLCTAHORLEALDAASGK 245
QY 110 LIMEHRROLPIATLNSF-----GI-----VANGYVA- 137
DB 246 EKMHYD--PELKTNESFOHTCGVSYHEAKAETASPEVADCPRIILPVNDGRILAI 302
QY 138 ----GSTCQ-----YSP-----FG 147
DB 303 NAENGKLCETFAANKGVNLQSNMPTDKPGLYEPTSPITITDKTIYVAGSVTDNFSTRETS 362
QY 148 CFVSGHDSATGEELMKNYFIPRA-----GEGDETCNDYEARMWTGANGQITDYDPT 200
DB 363 GVIKGFVNTGELLMA--FDGAKDPNAPISDEHTFTNS-----PNSWAPAAYDAKL 413
QY 201 NLVHGSTAVGPASETQRTGCTPGCLYGTNFAVRPTGETIWRHOTLPBDMDOECTPE 260
DB 414 DLVYLPKGVTTTPDIMGKNTPEQERYASSI-LALNATGKLAMSYQTVHNDLMDM---- 468
QY 261 MMTVNDVQPSSTEMEGLOSINPNAATGERVLTGVPCKTGTMMOFDAETGEFL----- 313
DB 469 ----LPAQPT-----LADITVN--GQKVPYIAPAKTGNIFFVLDNRGELVAPAPERK 515
QY 314 ---WARTNYQNMIESIDENGI-----VTNEDAILKEL--DVEYD----- 349
DB 516 VPGAAKGDVYTPQPSSELSFRPTKDLGADMGATMFDOLVCVRFHQMRKEGIFTTP 575
QY 350 ----VCPFLGGRDMPSAALNP----- 367
DB 576 SEQGLVFPNGLGMFEMGIGISVDPNREVALANPMALPFVSKLIPRGPNMEQPKDAKGT 635
QY 368 ---DSGI-----YFIPLNNVCYDMAVDOEFTSMQVNTSVTKLPBCKDMIGRIDAID 418
DB 636 GTEGSIQPOGYGVPTVTLN-----PFLS-----PGLPCCKQPMAGYISALD 676
QY 419 ISTGRLMSVERAANYS-----PV-----ISTGGVLFENGCT--DRFERAL 458
DB 677 LKTNEVVMKKRIGTPDQSMFPMPVPFPMNGMPLGISTAGNVLFIATADNLTARAY 736
QY 459 SOETGETLMOTRLATVASQA--ISYEVDGMQYVAI--AGGGSVYXS 501
DB 737 NMSGKELIMQGRLP--AGGQATPMTEYVNGKQYVVISAGHGSGFT 780

```

RESULT 11
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C/Accession: AG0523
R:Perkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrington, S.; Moutle, S.; O'Garra, P.
A:Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Paratyphi A
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0523
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <PAR>
A:Cross-references: GB:AL513882; PIDN:CAD01327.1; PID:Q16501455; GSPDB:GN00176
C:Genetics:
C:Gene: STY0191

Query Match	Score 327;	DB 2;	Length 796;
Best Local Similarity	12.1%;		
	Pred. No. 9.3e-16;		
	22.7%;		
	Missed matches 256;	Indels 240;	Gaps 31

Matches	157;	Conservative	66;	Matched	
OY	6	LWAS-----	AGALATLAAFAAQTVEPTDELLANPRAGEMISYGCONEYRHSPLTO	58	
Db	135	LWAGENDPOEINGTLRADATPA--ATSSSIDE-----	DMPATGRNGEGGRYSPLKO	185	
OY	59	ITTENGGOLALWYA-----	RGMPGKY--QVTPLIHDGVMYLANPGVYQIAIDKDTGD	109	
Db	186	ITANVHOLEAVEVFETGDKQPNDELTNEVPIKQGDITLYLCTAHORFLADLAASGR		245	
OY	110	LIMEHRROLPNIALTNSF-GIYANGV-----	IYA-----	137	
Db	246	EKMHHF--POLKTDSSFQHHYTCRGVSHYEAKAADTASFEVLAADCPRIILPVNDGRLLFAV		302	
OY	138	-----GSTCO-----	YSP-----	147	
Db	303	NAETGKLTETPANKGVLIQTNMDPTPGIYEFTSPILTDKTIYASVTDNFSTRETS		362	
OY	148	CFVSHGDSATGEELMRNTFIPRAGEBEDTEWGDYEFARMT--GAMGOITYDPVTNLYH		205	
Db	363	GVIRGADVNSGGLWMA--FDP--GAKPPTIPDDEHAFTFNSPNSWAAPAAKADIDLVL		418	
OY	206	GSTAVGPASETORCTPGGTLTYGNTFRAPRPDTEGELVWHRHQLPRNMDOECTFEPMATN		265	
Db	419	PMGVTPPDIWGNRRTPPEOERYASSI-LALNATGKLAWSIQYVHDLMDLPAOPTIAD		477	
OY	266	VDVQSTEMEGLOSTNPNAATGEREVLITGVCKTGIMQFDAETGE		312	
Db	478	ITVDGTV-----	PVIYAPAKGNIFVLDNRNGELVYPAPEKPPYGA	520	
OY	313	-----	LMADTYNQN-----	MIESIDENGIYV-VNEDA	339
Db	521	AKGVYAKTOPFSDLTFPRPKDLSGADIMGA--TMDOLYCVNWHQJRGIFTPPSEOG		579	
OY	340	ILKELDYEVYVCTPLGGRWPSAALNPDGSIYF-----	IFLANVCY-----	D	382
Db	580	TL-----	VFEQNLGMFEMGSIYVDPROVAIANPAAFLVSKLIPRPGNMEPPKO	631	
OY	383	MAAVDOETSMQVYNTS-NTY-----	KLPFGKMDIGRIDAIDISTGRTLMSVEPAAA	433	
Db	632	AKGTGTAGIGPOYGVPEVYTLNPFLLSPGGLPKCPQAMQYISALDLKNEIYMKKRIGTP		691	
OY	434	NYS-----	Y-----	LSTGGGVLFNGST-DRIYRALSOETGETINOTRLAT	473
Db	692	RDSMPPEPMPVPPFPMGMGLGPISTAGVAVFLIATADNYLRAIYMSNGEKLMDGRLP-		750	
OY	474	VASGA--ISYEVDSMQYVAL-AGGGSYS	501		
Db	751	AGGQATPMTYEVNKGQYVVISAGHSGFET	780		
RESULT	13				

S00943
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoace
 C:Species: Acinetobacter calcoacetius
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S00943
 R:Clifton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase
 A:Reference number: S00943; MUID:88289368
 A:Accession: S00943
 A:Molecule type: DNA
 A:Residues: 1-801 <CLF>
 A:Cross-references: EMBL:X07235; NID:938711; PIDN:CAA30222.1; PID:g38712
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Superfamily: respiratory chain
 C:Keywords: glucose dehydrogenase (pyrroloquinoline-quinone)
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:39-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91.93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 12.0%; Score 323.5; DB 1; Length 801;
 Best Local Similarity 21.8%; Pred. No. 1.7e-15;
 Matches 152; Conservative 75; Mismatches 193; Indels 277; Gaps 31;

20 PAFQVPTVDELNPARGENISTGQONERYRSPLOTITTEVGGQLQVMAKQPK 79
 152 PETAQAVPVAE-----SDMPAGRTQAGVRSYSPKQINDQVNDLKAAT--LRIGD 202
 80 V-----QVPLIHGVMYLANPGDVIOADAKTDGLWEHRRLPFIATLNSF- 127
 203 LKTDNDSEGTNQVPIKIGNMFCIAHQQLAIDPATGKERK--RDPMLKTKDSFQ 259
 128 ----- 127
 260 HLCRGVWYDANNTEFAISLSQKSSSTQCPKVFVFNVDRLVANAADGKACDFG 319
 128 -----GIYANG-VTVAGS--TCQSS--PGCCVSHDSATG 158
 320 ONGOVNIOEFPYAPGYNPTSPGIYTSYVAGSVTDNLSNKEPSG-VIRGYDVNTG 378
 159 EELRNATFIPRAGE-----EGDETWGNDYERAMTGMGOITDPVNLVHGSTAVGP 378
 379 KILW--VEDTGADPNAMRGEETTFVHNS-----PNMADPLADAKLIDV-VYPTGV- 427
 213 ASETQRTPGCLYG-----TNTFAVRPDTGIVRHOITLPRDMDDECTFEEM 262
 428 -----GTP--DIWGDRTLEKERYANSLMAINSTGLVWNPOTLHMDMDVPSOPS 479
 263 VTNVDVOPSTEMEGLOSINP--NAATGERRLTGV-----CTTG--TMMOPD 306
 480 LADIKKACQVPAIYVLTGKNAFVLDNR--NGQPIVPTKRPQVTKRGQKGEY 537
 307 AETGER-----LMA-----ROTNYQNMLESIDENGIYTVNED 338
 538 SKTQPSDINLAPQDKLTKDMMGATMLDQLMCRVSFKRLNTDGIYTPRESENGTL----- 592
 339 AILKELDEYDVCPFLFGGRMPSAALNDS-----GIYFI----- 374
 593 -----VFPGNLGVFEWKGMSVNDPROVAVNMPILGIFPVSRLLPADPNRAQAKG 641
 375 -----PLANNCYDMAAVDEFTSMADVNTSNVTKLPBGCDMIGKDAIDISTGRTLW 426
 642 AGTEGQVQPMGVY-----GVEISATLSPLGLPCQOPAMGVAGADLKTHEVW 691
 427 -----SVERAANYSPVLTSGGVLENCGT--DRYFRALSDQETGLM 467

Db 692 KKRIGTIROSLPRLPQPLAVKIGVPGCGSISTAGNWFVGATQDNYLRAFNVTNCKRLM 751
 QY 468 QTRLATVYASQA--ISYEVDGMQVVAI-AGGVSYGS 501
 Db 752 EARLP--AGGQATPMTYEINGQVYVIMAGSHSGFGR 786

RESULT 14

OPKEX
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda
 C:Species: Gluconobacter oxydans
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Clifton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 Mol. Gen. Genet. 229, 206-212, 1991
 A:Title: A single amino acid substitution changes the substrate specificity of quinop
 A:Reference number: S17716; MUID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CLF>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL Data Library, February 1992

A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, 'A', 214-808 <GOO>
 A:Cross-references: EMBL:X62710; NID:958416; PID:958417
 C:Genetics:
 A:Gene: gdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Superfamily: respiratory chain
 C:Keywords: glucose dehydrogenase (pyrroloquinoline-quinone)
 F:9-28/Domain: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:122-138/Domain: transmembrane #status predicted <TM5>
 F:91.93/Binding site: ubiquinone (Arg, Asp) #status predicted <TM5>
 F:470/Active site: Asp #status predicted

Query Match 10.9%; Score 295; DB 1; Length 808;
 Best Local Similarity 20.1%; Pred. No. 2e-13;
 Matches 141; Conservative 83; Mismatches 230; Indels 248; Gaps 25;

13 ALALLAFAQVPTVDELLA-----NPPAGEWISGQONERYRSPLOTIT 61
 128 AVIALFASLFTPHDISGLPTQIANASPADPDVPASEHAGKROACDRNSPINOINA 187
 62 ENVGQLOLVW-----ARGMQPKV--QVPLIHGVMYLANPGDVIOADAKTDGLIM 112
 188 TWVSNLKVAWHIHKDKMANSNDPQSTNATPIEFNNLTLMCSLHQKLEFVADGATGNVW 247
 113 EHRRL--PIATLNSFGI-----YANGVY----- 136
 248 VYDFRLQINPQFHLTCRGVSFHEPRTANMDSGNPAPTDCAKDSILFVNDGRLEVEDAD 307
 137 AGSTC-----QYSPGCGV----- 150
 308 TGTCTSGSGNNGEIDLRYVNPQPTTPGQLEPSPVITDKLILANSATITNGSVKQASGA 367
 151 -SGHDSATGEELM-----RNYFTIPRAGEGDETWGNDYERAMTGMGOITDPVNLVHY 205
 368 TQAFDYVTGKRVVFDASNDPQPLPDSHPVFPNPSNHYIS-----SYDANLNLVYI 422
 206 GSTAVGPASETQGRFP-----GFTLYGTNTRF-----AVRPDTGIVRHOITLPRDMDDECT 257
 423 PMGV-----GTPDQMGDRKTDSERFAPGIVALNADTGLAMFVQTVHNDLMDHEL 473

Fri May 24 11:27:30 2002

walick-934-128.rep.rpt

258 TREMNATVNDVQSTEMEGLOSINPNATGRRVNLGVPCKTGTMMQDAETG-ETLNAR 316
Db 474 PSQPSLVDTOKDGLVPAIYA-----PRTKIDIFVLDRTGKEIYAP 517
QY 317 DT-----NYOMIESIDENGIVTVNEDALKEEDV-----EYD 349
Db 518 ETVPQAGABDHTSPQPMQO-----LTLRKNPLNDSDIMGCTIFQDMCSIFITLRE 574
QY 350 -----VCPFLGGRDMSAALNPDGIFIFPLNNVCIDMAAYDO-----388
Db 575 GFPTPSLKSLIFPDIDLMFEMGLAVDPQVAFANPISLPVSQLVPRPGNPLME 634
QY 389 -----EFTSMQVNTSVNTKLPKQDMT-----GRIDADISTGRTL 425
Db 635 ENAKGTGEGTLOHNTGIFPVNLFHPLDVLPLPGCIKPCRRPPWGIYAGIDLKTRKV 694
QY 426 WSYERAANYS-----PV-----ISTGGVLF-NGSTDRTYPRALSOETG 463
Db 695 WOHNRGTLRDSMVGSSLPPLPIKIGVPSLGSLTAGNLGELTASMDYIIRAYMLTGT 754
QY 464 ETIMOTRLATVASGALSYVDGMOYVAIAGGVSYSGSLNS 505
Db 755 KVLMDRLPAGQANPTITAIKQYI-----VYAGGINS 790

RESULT 15

F83360
glucose dehydrogenase PA2290 [Imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence: Revision 15-Sep-2000 #text-change 31-Dec-2000
C:Accession: F83360
C:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Watrener, P.; Hickey, M.J.; B.
Kistover, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; Lm.
adman, S.; Olson, M.V.
.. Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:2043737
A:Accession: F83360
A:Status: preliminary
A:Residues: 1-803 <STO>
A:Cross-references: GB:AE004654; GB:AE004091; NID:99948311; PIDN:AG05678.1; GSPDB:GN001
A:Experimental source: strain PA01
A:Genetics: PA2290
C:gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 10.1% Score 273.5; DB 2; Length 803;
Best Local Similarity 21.5% Pred. No. 7.5e-12; Indels 285; Gaps 32;
Matches 158; Conservative 69; Mismatches 222; -----36
QY 6 LHM-----ASKGALLAPAFNAVTPYTDLLANP-----O 81
Db 103 LLMLEPFRRLADGAPPLGTALGAVAVLAGAAVGSQ-FTNPGQIAGRDRDSGMTST 161
QY 37 -----PAGEMISYGONQENTRHSPITQITTEVNGQLVWARGMOPGV-----219
Db 162 APAMPDGMQAGRTREEDRISPKOITPANVGOLEARN-IRTDLEPTADDDLEFNE 219
QY 82 VTPLHDGVWYLANPGDVIOAIDAKTGDILWEHRPLNATLNSFGIANGVYASGC 141
Db 220 NTPLVKMGMLYACTASKYVALDPPDTCALMREPDQIOSPVGFKG-----AHMTC 270
QY 142 -----QY-----SPFG-----CFVSGHDSATGEELMRYPTPRAGE-----173
Db 271 RGVSYDEQIARSDVAGAPPALSEAGKAVASCPRLFLPTADALTAIINADNGKVCED 330
QY 174 -----GFTWGNDE-----ARMTCAMQOITYDPVNL-----VHY 205
Db 331 FGKAVDLTAGIGFTPGGYSTSPAAVTRNLVILGHTVDNESTNEPSCVIRADVDH 390
QY 206 GSTA-----VGPASRTQRTG-----GTLY-----GTN 229

Db 391 GLKLVNMNDGNDPDETEBLAPGKFTYRNSPNMWSLASVDEKLGQVLYPLGNQMPDONGNR 450
QY 230 T-----RF-----AVPDTGELVWRHOTLPROMNDOECTFEMNTVNDVQSTEMEGLOST 494
Db 451 TPGAERFSAGVALDLNLTGKLRMYOFTPHDLMDMD-----VQSPTL-----310
QY 281 NPNATGERRYLTGVPCKTGTMMQDAETG-----ROTNYOMIESIDENGIVTVNEDALKEEDV 348
Db 495 DKTADGVKPALI-APFKQGSILYVLDRODGPPIYPIREVPAGABEDHTAPQARS 553
QY 311 -----EFLMA-----ROTNYOMIESIDENGIVTVNEDALKEEDV 598
Db 554 NLRLPPLTERDMMGSSPPDDMLCRIOFRSLRKEGQYTPSPQSL-----SNDVYNTSVTK--403
QY 349 DYCEPFLGGRDMSAALNPDGIFIFPLNNVCIDMAAYDOET-----SNDVYNTSVTK--403
Db 599 -IYPONVGVFMVGVSVDVRLDFTSPNYAFVSOMVPRDVPSCSKREGESGVQPM 657
QY 404 -----LPGKDMIGRIDADISTGRTLMSEVERAA-NISPV-----438
Db 658 GAPVYIMHPFMSPIGLPCQAPSMGVDAGIDLTITAKYVMOHKNKGTSRDNTVPYIGLTVGV 717
QY 439 -----ISTGGVLFNGGT-DRYPRALSOETGETIMOTRLATVASGA--ISYE-VDMQ 488
Db 718 PSMGGSITTAGVAFSLGTLDDYLRAYDKRQOLMQARLP--AGGATPMSTYTGKDG 775
QY 489 YVAT-AGGVSYS 501
Db 776 YVLVAGHGHSFCT 789

Search completed: May 24, 2002, 10:21:08
Job time: 346 sec

Fri May 24 11:27:30 2002

wallick-934-128.pcp.rpr

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:52 ; Search time 28.06 Seconds
(without alignments)
696.842 Million cell updates/sec

WALICK-934-128.PSP

Title: 2705
Sequence: 1 MKRPSLMAAGALALAAP.....GMOYVIAAGGVSYGSLINS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483.5	17.9	739	1	DHET_ACEBU
2	479.5	17.7	738	1	DHET_ACEBU
3	477.5	17.7	742	1	DHET_ACEBU
4	466.5	17.2	757	1	DHET_GLUOX
5	435.5	16.1	600	1	XOXF_PARD
6	434.5	16.0	626	1	DHMI_METEX
7	434.5	16.0	631	1	DHMI_METEX
8	422.5	15.6	631	1	DHMI_PARDE
9	417.5	15.4	631	1	EXXA_PSEAE
10	393.5	14.5	571	1	DHMI_METME
11	342.5	12.6	796	1	DHGA_ECOLI
12	323.5	12.0	801	1	DHGA_ACTICA
13	295.5	10.9	808	1	DHGA_GLUOX
14	293.5	9.4	809	1	QUIA_ACTICA
15	231.5	8.6	790	1	QUIA_XANCU
16	128.5	4.7	392	1	YEGU_ECOLI
17	126.5	4.7	827	1	AFSK_STRCO
18	118.5	4.4	799	1	AFSK_STRCO
19	116.5	4.3	593	1	SPG2_STRSP
20	115.5	4.3	353	1	YXAL_BACSU
21	115.5	4.3	353	1	CADN_MOUSE
22	115.5	4.3	443	1	PORD_PSEAE
23	115.5	4.3	954	1	FLEY_CADCR
24	114.5	4.2	415	1	Y232_RICPR
25	114.5	4.2	796	1	COPP_SCHPO
26	110.5	4.1	621	1	ASPA_AERSA
27	110.5	4.1	3317	1	CADN_RAT
28	108.5	4.0	3354	1	CADN_HUMAN
29	106.5	3.9	1276	1	SCAP_CRIGR
30	106.5	3.9	966	1	PHSI_SOLTU
31	105.5	3.9	1355	1	RPOD_ANASP
32	105.5	3.9	697	1	NANB_STRPN
33	103.5	3.8	649	1	ACES_DROME

34	103.5	3.8	974	1	PHS2_SOLTU
35	103.5	3.8	1365	1	GFPS_STRDO
36	102	3.8	595	1	YER5_SCHPO
37	102	3.8	790	1	PLMN_PIG
38	102	3.8	1323	1	RPO_SMTXA
39	102	3.8	1645	1	OMPB_RICTY
40	101.5	3.8	1637	1	MKSP-STAU
41	101	3.7	894	1	FOX2_NEUCR
42	101	3.7	1012	1	POLG_IBDVO
43	100	3.7	466	1	MMOB_RAT
44	99.5	3.7	807	1	AFSK_STRGR
45	99.5	3.7	3567	1	ERY2_SACR

ALIGNMENTS

RESULT	ID	Score	Query Match	Description
1	DHET_ACEBU	483.5	17.9%	739 AA
AC	044002: 007952;			STANDARD;
DT	15-DEC-1998 (Rel. 37, Created)			PRT;
DT	15-DEC-1998 (Rel. 37, Last sequence update)			739 AA.
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).			
GN	ADH.			
OS	Acetobacter europaeus.			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;			
OC	Glucosyltransferase.			
OX	NCBI_TaxID=33995;			
RN	SEQUENCE FROM N.A.			
RA	STRATIN-DES11 / DSM 6160;			
RC	Turner C.A.K.; to the EMBL/GenBank/DBJ databases.			
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.			
CC	-1- COFACTOR: POO AND HEME (BY SIMILARITY).			
CC	-1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, AND TWO SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL: X82894; CAA58066.1; -			
CC	EMBL: Y09480; CAA70688.1; -			
CC	HSSP: Q92477; 1FUG.			
DR	InterPro: IPR001479; Bac_POO.			
DR	InterPro: IPR002372; Bac_POO_repeat.			
DR	InterPro: IPR000345; CytC_heme_bind.			
DR	Pfam: PF01011; Bacterial_POO_6.			
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C_1.			
DR	Oxidoreductase; POO; Heme; periplasmic; Membrane; Signal.			
KM	SIGNAL			
FT	CHAIN			
FT	BINDING			
FT	IRON (HEME AXIAL LIGAND) (BY SIMILARITY).			
FT	IRON (HEME AXIAL LIGAND) (BY SIMILARITY).			
FT	SEQUENCE			
SO	SEQUENCE			

Query Match 17.9%; Score 483.5; DB 1; Length 739;

Best Local Similarity 26.2%; Pred. No. 1.3e-28;
Matches 153; Conservative 77; Mismatches 237; Indels 117; Gaps 16;

```

QY 10 SAGALLALLAPAPAQVPTVDE-----LLANPAGEMISYCONENYHSPLOTIT 61
DB 17 TAGTICALLISGATAMASADDGATGATGAIHADHPGMMYMTGTYEQRTSPLODQINR 76
QY 62 ENNGOOLYVARKMOPGVQV--TPLIHGVMYLANPGVIAIDAKTGDLWEHRRLP- 119
DB 77 SNGNKLAMVYLDLDTNRGQEGPLVDGVMTATNMSMKRAVDAATGKLMSYDPVPG 136
QY 120 NIA-----TLNS-----FGI----- 129
DB 137 NIADKGCCTVNRGAAYNNGKYVETGDFGRILALDAKTGLVMSVNTIPPEALGKORSY 196
QY 130 -----VANGYIVAGSTCOYSPFGC--FVSGHDSATGEELMANYFIPRAGEGDETCGN 180
DB 197 TYDGAAPRIAKGRVITIGN--GGSEFGARGFYAPFAETGKVDMEFTAPNPKNEPHTASD 254
QY 181 D-----YEARMTGA-----NGQITDPVTNLVHYGSTAVGPASETORGTGGT 224
DB 255 SVLANKAVQWSPGAMTRGCGGTWDSIVDPVADLVYLGNGSPNNYRSEBGKD 314
QY 225 LYGTRAVRVPDTEGLVYHQTLPKDMDOECTFEAMVTNVDQGSTMEGLSINPNA 284
DB 315 NLFGSIVALKPTEGEVYVHFQETPMQDMFTSVQOIMTLDPINGET----- 362
QY 285 ATGERRVLTGVPCKTGTMMQFDETEGELMADNTYONMIESID-ENGIVTNEADILKE 343
DB 363 -----RHVYHAP--KNGFYIIDAKTGEFTSGKNYVVMASGLDKTRIPNDALYTL 417
QY 344 LDVEYDCPTFLGRRMPSAALNPDGSIYFIPLVNVCYDMAVDOEFT-SMDVYNTS--- 399
DB 418 TGEKMGITPDDLGHNFAAMAFSPKTYLVIYPAQOVPLTYNOVGSTPHPDSNGLDLM 477
QY 400 NTKKLPPG-----KDMIGRIDIDISTGRTLSVERAANTSPVLSGGGVLENGST 451
DB 478 NKVGIDSPSEAKQAFYKDKLKWIVAMDPOKAEAMRVDRHKGPMWGGILATGDLPGOLA 537
QY 452 DRYFRALSGEGETLMQRIATVASGAISYEVDGMOYVIAAG 495
DB 538 NEFHAYDATNGSDLHFRPADSGIAPVYTYLANGKQYVAVEVG 581

RESULT 2
DHET_ACEPO
ID DHET_ACEPO STANDARD; PRT; 738 AA.
AC P28036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (Ec 1.1.99.8).
OS ADA.
NC Acetobacter polyoxogenes.
CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
CC Acetobacter.
OX NCBI_TaxID=439;
[1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NB11028;
RX MEDLINE=91159482; PubMed=2001402;
RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
RT Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.,
RT "Cloning and sequencing of the gene cluster encoding two subunits of
RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes."
RL Biochem. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.

```

```

CC CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC CC SPACE (POTENTIAL).
CC CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb.ch/announce/
CC CC or send an email to license@isb.sib.ch).
CC CC -----
CC DR EMBL: D00635; BAA0528.1;
CC DR PIR: S14270; S14270.
CC DR HSRP: Q924J7; S14270.
CC DR InterPro: IPR001479; Bac_POO.
CC DR InterPro: IPR002372; Bac_POO_repeat.
CC DR Pfam: PF01011; Bacterial_POO.
CC DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
CC DR PROSITE: PS00190; CYTOCHROME_C; 1.
CC KM Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
CC FT SIGNAL 1 35
CC FT CHAIN 36 738
CC FT BINDING 650 650
CC FT BINDING 653 653
CC FT METAL 654 654
CC FT METAL 654 654
CC SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

```

Query Match 17.7%; Score 479; DB 1; Length 738;
Best Local Similarity 25.9%; Pred. No. 2.8e-28;
Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

```

QY 10 SAGALLALLAPAPAQVPTVDE-----LLANPAGEMISYCONENYHSPLOTIT 61
DB 17 TAGTICALLISGATAMASADDGATGATGAIHADHPGMMYMTGTYEQRTSPLODQINR 76
QY 62 ENNGOOLYVARKMOPGVQV--TPLIHGVMYLANPGVIAIDAKTGDLWEHRRLP- 119
DB 77 SNGNKLAMVYLDLDTNRGQEGPLVDGVMTATNMSMKRAVDAATGKLMSYDPVPG 136
QY 120 NIA-----TLNS-----FGI----- 129
DB 137 NIADKGCCTVNRGAAYNNGKYVETGDFGRILALDAKTGLVMSVNTIPPEALGKORSY 196
QY 130 -----VANGYIVAGSTCOYSPFGC--FVSGHDSATGEELMANYFIPRAGEGDETCGN 180
DB 197 TYDGAAPRIAKGRVITIGN--GGSEFGARGFYAPFAETGKVDMEFTAPNPKNEPHTASD 254
QY 176 -----ETWGNDEYARMTGA-----NGQITDPVTNLVHYGSTAVGPASETOR 218
DB 255 SVLANKAVQWSPGAMTRGCGGTWDSIVDPVADLVYLGNGSPNNYRSEBGKD 314
QY 219 GTPGTLVGTNTRAVRVPDTEGLVYHQTLPKDMDOECTFEAMVTNVDQGSTMEGLS 307
DB 308 SEGKNDLVLSGIVALKPTEGEVYVHFQETPMQDMFTSVQOIMTLDPINGET----- 357
QY 279 SINPNAATGERRVLTGVPCKTGTMMQFDETEGELMADNTYONMIESID-ENGIVTNE 337
DB 358 -----NGETRVYIYHARKNGFYIIDAKTGEFTSGKNYVVMASGLDKTRIPNDALYTL 417
QY 338 DATLKELDVYDCPTFLGRRMPSAALNPDGSIYFIPLVNVCYDMAVDOEFT-SMDVY 396
DB 411 DALYTLTGKEMWGITPDDLGHNFAAMAFSPKTYLVIYPAQOVPLTYNOVGSTPHPDSM 470
QY 397 NTS--NTKKLPPG-----KDMIGRIDIDISTGRTLSVERAANTSPVLSGGGV 445
DB 471 NLGLDNKAVGIDSPSEAKQAFYKDKLKWIVAMDPOKAEAMRVDRHKGPMWGGILATGDL 530
QY 446 LKNGTDYFRALSGEGETLMQRIATVASGAISYEVDGMOYVIAAG 495

```

Db 531 LFGGLANGERHAYDATNGSDLFHFAADSGIIPVYTIANGKOYAVEVG 580

```

RESULT 3
DHET_ACEAC STANDARD; PRT; 742 AA.
ID DHET_ACEAC
AC P18278;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADHI.
OS Acetobacter aceti.
OC Acetobacter; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OC NCBI_TaxID=435;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RX MEDLINE=89255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=95289964; PubMed=7772016;
RA Gglier G.E., Gilles I.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens."
RL Biochem. J. 308:375-379(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.
CC -1- SUBUNIT: TETRAMER OF NON IDENTICAL CHAINS (DEHYDROGENASE,
CC CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90004; BAA14058.1;
CC DR PIR; J50326; J50326.
CC DR HSSP; 092437; 1FLG.
CC DR InterPro; IPR001479; Bac_POQ.
CC DR InterPro; IPR002372; Bac_POQ_repeat.
CC DR InterPro; IPR000345; CYC_heme_bind.
CC DR Pfam; PF01011; Bacterial_POQ_6.
CC DR PROSITE; PS00363; BACTERIAL_POQ_1; 1.
CC DR PROSITE; PS00364; BACTERIAL_POQ_2; 1.
CC DR PROSITE; PS00190; CYTOCHROME_C_1.
CC DR OXidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
CC KW SIGNAL
CC FT CHAIN 1 35
CC FT ACT SITE 36 742 ALCOHOL DEHYDROGENASE [ACCEPTOR].
CC FT BINDING 343 343 BASE (POTENTIAL) (BY SIMILARITY).
CC FT BINDING 649 649 HEME (COVALENT) (BY SIMILARITY).
CC FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
CC FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SEQUENCE 742 AA; 81521 MW; 9C6C9268DAB825A CRC64;

```

Query Match 17.7%; Score 477.5; DB 1; Length 742;
 Best Local Similarity 26.2%; Pred. No. 3.7e-28;
 Matches 161; Conservative 89; Mismatches 202; Indels 163; Gaps 24;

```

QY 11 AGALLAARFAVPTVPEDEL--ANPPAGEWISYGOENYRHSPLTOTTEVGO 68
DB 23 AALPYAVAPRADGCGNTGEIALLHADDPENMLSYGRTYSQRYSPDQIRSNVGLK 82
QY 69 LVMARGMOPGRVQ--VTPLIHDGYVYLANPGDVLOAIDATKGLDIMEHRQLP-NIA---- 122
DB 83 LIGYTLDTNRGOEATPLVVDGIMVATTNMSKMEALDAATGKLIMQYDPKYGNIADKGC 142
QY 123 --TLN-----SEGI----- 129
DB 143 CDTVRKAGQWNGKVFCTFGRLVLAADATKGGKVMVNTIPADASIGKORSTYDGAVR 202
QY 130 VANGVIVASGTCQYSPFC--FVSGHDSATGEELMRVYFIPRAGEGD----- 175
DB 203 VAKGLVLIJN--GGAFFGARFVSFAFDETKLKRFTYVPPNKNKNEPDHAASNNILMKA 260
QY 176 -ETWGDYEARMMT-----GANGQITPDVNTLVHYGSTAVGAPASETORPGTLY---- 226
DB 261 YKTWGP--KGAWRGQGGGTWDSLYDPVSLTY--LAVG-----NSPMNYKTRSE 309
QY 227 --GN-----TRFVRPDTEIWRHQTLPKRWDOECFFEMMYTINVDPSTEMEGLOSI 280
DB 310 GIGSNLFGISVALKPEETGEYWHFQATPMDQDYTSVOQITLDMVK----- 358
QY 281 NPNATGE--RRVLGVPCKTGTWMOFDETGELWADRTNQNLESID-----EN 330
DB 359 -----GEMRHVIVHAF--KNGFFYVDATKTGELSGKNVYONMANGDPLTGRPMYNDP 411
QY 331 GIVTNEADAILKELDEYDVCPTFLGGRDMPASALNPDSGIFIFILNNVCY----- 381
DB 412 GLYLING-----KFWYGI-PGPLAHNEMAMAYSPKTHLVYIIRAHQIPFYKNQVGF 463
QY 382 ----DMMAVDDEFTSMYVNTSNTKLPKGDMDIGRIDAIDISTGRTMSVRAAANTSP 437
DB 464 KPHADSMWVNGIDMTKNGLPDPE--ARFAYIKDLGMLLADVPKMTVWKIDHKGPWNG 522
QY 438 VLSTGGVLFNGTDRYFRALSOETGLMOTRLATVYASQALSYEVDDMOYVA----- 491
DB 523 ILATGDDLFOGLANGERHAYDATNGSDLYKFKDAQSGIIPVYTIANGKOYAVEVG 582
QY 492 ----TAGGVSYSG 502
DB 583 GIYPISMGGVGRTS 597

```

```

RESULT 4
DHET_GLUOX STANDARD; PRT; 757 AA.
ID DHET_GLUOX
AC 005342;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH
DE subunit I).
GN ADHA.
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconobacter.
OC NCBI_TaxID=442;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RC STRAIN=IFO 12528;
RC MEDLINE=97208225; PubMed=9055427;
RA Kondo K., Horinouchi S.;
RT "Characterization of the genes encoding the three-component membrane-
RT bound alcohol dehydrogenase from Gluconobacter suboxydans and their
RT expression in Acetobacter pasteurianus."
RL Appl. Environ. Microbiol. 63:1131-1138(1997).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL)
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.

```

```

CC -1 SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1 SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERILASMIC
CC SPACE (POTENTIAL).
CC -1 SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D86375; BAI19753.1; -.
CC HSSP: Q924J7; 1F1G.
CC InterPro: IPR001479; Bac_POO.
CC InterPro: IPR002372; Bac_POO_repeat.
CC Pfam: PF01011; Bacterial_POO.
CC PROSITE: PS00363; BACTERIAL_POO_6.
CC PROSITE: PS00364; BACTERIAL_POO_1; 1.
CC PROSITE: PS00190; CYTOCHROME_C; 1.
CC Oxidoreductase; P00; Heme; Periplasmic; Membrane; Signal.
CC SIGNAL 1 34
CC CHAIN 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
CC MOD_RES 35 35 PYRROLOIDONE CARBOXYLIC ACID.
CC ACT_SITE 342 342 BASE (POTENTIAL).
CC BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
CC BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
CC METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;
SO

```

Query Match 17.2%; Score 466; DB 1; Length 757;
 Best Local Similarity 25.0%; Pred. No. 2.8e-27;
 Matches 153; Conservative 89; Mismatches 227; Indels 144; Gaps 19;

```

OY 7 LMSAGLALIAA--PAFAQV---TPVDELLANPPAGEWMSYQONENYHSHSLTQIT 61
DB 16 LLSCAALAFSAAPVAFQEDTGTALTSSDNGHP--GDWLSYGRSISEQYSPDLQINT 74
OY 62 ENVGOLQVWARGMOPGKQVQV--TPLIHGVWYLANPGDVIQAIQDAKTDLIWEHRQLP 119
DB 75 ENVGKTLAMHVDLDTNRQSEPTPLIVGVWYATTNMSKMLDAATGKLLMYDPKVG 134
OY 120 NTA----- 122
DB 135 NIADNGCCDVTYSGAAYWNGKVFETPDRLIALDAKTGLVMSVYTIPEKQOLGHORSY 194
OY 123 TLNSEGIYANGVIYAGSCQYSPGCG--FVSGHDSATGEELMKNYFIPRAGEGD----- 175
DB 195 TVDGAPRIAKKRVLIGN--GGAEGANGFVSAPDAETSKIDMFFVVPENKPPGGAASD 252
OY 176 -----ETWGNDEYEARMMTG--AWGQITDPTVNLVHYSGTANVPASEFQRTPGGT 224
DB 253 DILMSKAVPTWKGKNGAKWQGGGTWDSLYVDPVTDIYLVGVNGSPNNYKFEKGGCD 312
OY 225 LYGTNFEAVAPDGEIYVHRQTLPRDNDQSETEPMATNVADYOPSTEMEGLASINPNA 284
DB 313 NLFGLSVIALNPDTGKVMHFQETPMDEWDYSVOOIMTLDMPV----- 356
OY 285 ATGE-RKVLTVGPCKTGTMMQDAETGEFLMARDNTYOMIESIDE-NGIYTVEDAILK 342
DB 357 -NEMRHVYIHAH-KNGFEYIIDAKTGKITKPTTYENMANGDDPYVGRPNVYDALMT 414
OY 343 ELDEYDVCTPLFGRDMSALNDSCGYIPLNNV-----CYDMAVADOE 389
DB 415 LTGKPHLIGTPELGHNFAAMAYSPKTKLYIYPAQIQLILDGQKGFKAIVHAMNGLD 474
OY 390 FTSMDYNTSNVTKLPGCKMIGRIDAIDISTGRL--MSVRAANLS-----PVL 439
DB 475 MNKIGLFDNDDEHVAAKKDF-----LKYVKGWTVAMDPEKMAPATTINHKGPWNGGL 528

```

```

OY 440 STGGVLENGGDRYFRALSGEETIMQTRLATVASGCAISIEVDGMQYAI----- 492
DB 529 ATAGNVITQIANGFENFAYIATNTGNDLYSPPAOSAIAPVYTTANGKQYAVEVGWGI 588
OY 493 ---AGGVSYSG 502
DB 569 YPFLYGVARTSG 601

```

RESULT 5
 ID XOXF_PARDE STANDARD; PRT; 600 AA.
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_Taxid=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harms N.;
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 380-600 FROM N.A.
 RC STRAIN=PD 1235;
 RA MEDLINE=92041583; PubMed=1657873;
 RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Olmann L.F.,
 RA Stoutamer A.H.;
 RT "Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c553 of Paracoccus denitrificans and characterization of
 RT the mutant strain.";
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1 COFACTOR: P00 (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U33436; AAC44555.1; -.
 CC EMBL: M75583; AAA25574.1; -.
 CC DR PIR: A41378; A41378.
 CC HSSP: P38539; 4AAH.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC Oxidoreductase; P00; Signal.
 CC SIGNAL 1 21
 CC CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.
 CC ACT_SITE 318 318 BASE (POTENTIAL).
 CC SEQUENCE 600 AA; 65159 MW; DCA96F1BC5A3CE CRC64;

Query Match 16.1%; Score 435; DB 1; Length 600;
 Best Local Similarity 25.7%; Pred. No. 4.3e-25;
 Matches 154; Conservative 80; Mismatches 227; Indels 138; Gaps 23;

```

OY 10 SAGALALIAAPFAFAQVPTVDELLANP-----PAGEWISYQONENYRSPPLTQITTE 62
DB 6 NGACTALLMSGTA-----LANEQRAGDRQAPQMAIQMGYANTRTSTLDQINKD 56
OY 63 NVGOLQVW--ANGMOPGVQVYPTLIHGVWYLANP-GDVIQAIQDAK-TGDLIWEHR-RQ 117
DB 57 NVKDLRVAMTFSTGVLRGH-EGSPVLIQDVWYVHPPFNRRFALDLNDNGKILMYRPEQO 115

```

RA	Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;	
RT	The refined structure of the quinoprotein methanol dehydrogenase	
RL	from <i>Methylobacterium extorquens</i> at 1.94 Å.	
CC	Structure 3:177-187(1995).	
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +	
CC	reduced acceptor.	
CC	-1- COFACTOR: POQ.	
CC	-1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION	
CC	OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.	
CC	-1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER	
CC	MEMBRANE.	
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL collection -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
CC	DR EMBL: M31108; AAA25380.1; -	
CC	DR PIR: S07908; S07908.	
CC	DR PIR: J00706; J00706.	
CC	DR HSR: P38539; 4AAB.	
CC	DR InterPro: IPR001479; Bac_PQO.	
CC	DR InterPro: IPR002372; Bac_PQO_repeat.	
CC	DR Pfam: PF01011; Bacterial_PQO_7.	
CC	DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.	
CC	DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.	
CC	CC OXIGENOLACTASE; POQ; Signal; Methanol utilization; Periplasmic.	
KM	Signal: Methanol utilization; Periplasmic.	
FT	CHAIN 1 27	
FT	CHAIN 28 626	
FT	DISULFID 130 131	
FT	DISULFID 413 442	
FT	ACT_SITE 330 330	
FT	ACT_SITE 626 AA: 68434 MM: 64988DDAFAFDAD34C CRC64;	
FT	SEQUENCE	
CC	-----	
CC	BASE (POTENTIAL):	
CC	-----	
CC	METHANOL DEHYDROGENASE SUBUNIT 1.	

```

Query Match Similarity      16.0% ; Score 4.93 ; Db 17 ; Length 226
Best Local Similarity      24.5% ; Pred. No. 5,4e-23 ;
Matches 148 ; Conservative 91 ; Mismatches 208 ; Indels 158 ; Gaps

```

Db 386 KSVDLKTGQPVDPPEYGRMDHLAKDIPASMGYHNGDSYDKELEFFMGJNHICMDW 445
 QY 384 MAVDEFTSMQVYNSNTKLPKPK-----DMIGRIDAIDISTGRTLSVERAAANYSP 437
 Db 446 EPEMLPYRAGQFFVGATLMYTPGPKGRONTEGQIKAYNAITGDYKKEKMERFAVMWG 505
 QY 438 VLSGGVLFNGGTDYRFRALSOETGTLWOTRLATVASGOATSYEDVMQVVAIAGGV 497
 Db 506 TMAATGDLVFGTLDGLYKARDSDTGLMKFKIPSGAIGYPMYTHKGTQYVAI----- 560
 QY 498 SYGSG 502
 Db 561 YYGVG 565

RESULT 7
 DHM1_METOR
 ID DHM1_METOR STANDARD: PRT: 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MEDH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
 RC STRAIN-ATCC 27886 / DSM 760 / NCIB 11278;
 RX MEDLINE=8908094; PubMed=2459109;
 RA Machin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 Methylobacterium organophilum XX methanol dehydrogenase structural
 gene".
 RT J. Bacteriol. 170:4739-4747(1988).
 RL 3.
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: M22629; AAA50289.1;
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 28
 FT CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131 BY SIMILARITY.
 FT DISULFID 413 442 BY SIMILARITY.
 FT ACET_SITE 330 330 BASE (POTENTIAL).
 FT SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DFF CRC64;

Query Match 16.0%; Score 434; DB 1; Length 626;
 Best Local Similarity 24.5%; Pctd. No. 5.4e-25;

Matches 148; Conservative 91; Mismatches 208; Indels 158; Gaps 23;
 QY 10 SAGALALLA-AP-APAOYTPYDELAPNPPAGE-MISVGNQOEYRSPITQITENQO 66
 Db 7 SVSALMLALAPALSSVAYANDKIELSKDDMMVWPKAYDSNNSELKOVNKSVMQ 66
 QY 67 LQLVW-ANGNPGKVVQYTPPLHSGVMY-----LANPGDY----- 99
 Db 67 LRPAMTFSTGLNGH-ESAPLVVDKMYVHTSPFNPTFALDLDPCGILLMOQPKQNPAA 125
 QY 100 -----IOAIDKGTGLMEHRROLPIATLN 125
 Db 126 RAVACCDLVNRLAWPDPGKTPALLIKTQDLDRHYVALNAEIGETVWK----- 173
 QY 126 SFGIVANGVIVAGSTCCQSPF-----GC-----FVSGHDSATGEELMNFYI- 167
 Db 174 -----VENSIDIKVSGSTLIAPYVKKVYIIGSGAELGVGYLAVDVKTGGQVWRATNG 229
 QY 168 -----PRAGEE--GDETGNDEYARMNTGA-----KQITYPYTNLVHYS 207
 Db 230 PDKDLLADDENVKNANHYGKGGTATWBD--AKKIGGTMWGVAYDPGTMLYFGT 286
 QY 208 TAVPASETORGPFGTLYGTNTREAVRDPGTGEIYWRHQITLPRDMMDOCTEFEMVTVND 267
 Db 287 GNPAPMNETMK--PGDKW-TMTIFGRDADTGEAKFGYOKTPHDEMVDYAGVNMW----- 338
 QY 268 VQSTEMEGLOSINPNAATGERVLTGVCPTKTMQFDATGEPLAR--DTNYQNT 324
 Db 339 --PSEKQD-----KDKTRKLLTHDRNGIYITLDRDGLVSAANKLDDT--VNVF 385
 QY 325 ESIDENQIVYVNDALIKELD-VEVDVCPTEFLGGRDMPASALNPDGSIYFPLNNVYDM 383
 Db 386 KTYVDLTGQPVDPPEYGRMDHLAKDIPASMGYHNGDSYDKELEFFMGJNHICMDW 445
 QY 384 MAVDEFTSMQVYNSNTKLPKPK-----DMIGRIDAIDISTGRTLSVERAAANYSP 437
 Db 446 EPEMLPYRAGQFFVGATLMYTPGPKGRONTEGQIKAYNAITGDYKKEKMERFAVMWG 505
 QY 438 VLSGGVLFNGGTDYRFRALSOETGTLWOTRLATVASGOATSYEDVMQVVAIAGGV 497
 Db 506 TMAATGDLVFGTLDGLYKARDSDTGLMKFKIPSGAIGYPMYTHKGTQYVAI----- 560
 QY 498 SYGSG 502
 Db 561 YYGVG 565

RESULT 8
 DHM1_PARDE
 ID DHM1_PARDE STANDARD: PRT: 631 AA.
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MEDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RC MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 structural gene from Paracoccus denitrificans".
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH

ON METHANOL (IN P.DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE TOTAL CELL PROTEIN).
 -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: M17339; AAA88366.1; -
 HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KM Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 135 136 BY SIMILARITY.
 FT DISULFID 418 447 BY SIMILARITY.
 FT ACT_SITE 335 335 BASE (POTENTIAL).
 FT SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;

 Query Match 15.6%; Score 422.5; DB 1; Length 631;
 Best Local Similarity 24.8%; Pred. No. 4e-24; Indels 179; Gaps 26;
 Matches 153; Conservative 85; Mismatches 201;

 QY 4 TSLMASAGLALL-AAPAFQVPTVDL--LANPAGWISYGCNENTRHSPLQIT 60
 DB 12 TSLMAAAMGLAVITTPARA-----NDQLVELAKDPA-NMWTGRDYNAONTSEMDIN 65
 QY 61 TENYGOLOLV--ARGN-----OPKV-QVTP- 84
 DB 66 KENKQALPAMSFSTGVIAHNGSTPLVGDPMFHPPTNTFPALDNEPGLIMQNKPK 125
 QY 85 -----LIHGVNLYANPGD-----VIOADAKTGDLIMHRRQL 118
 DB 126 QNPATVACCDVVRGLIAY--PGDQVYKPLFTQDGHIVAMDAETGETRW- 177
 QY 119 PNIATLNSFGIYANGVIYASTCOYSP-----GC-----FSGHDSATGEEL 161
 DB 178 -----IMENSDIKVGSITLTITARYIKDLVYSSGAELGVYGYTAVDKSGEMR 227
 QY 162 WKNYFI-----PRAGEE--GDETGNDYARWMTGA--WQITDPVT 200
 DB 228 WRFATGPBEELLADENAPRPHYGNKLGLETWEGD--AMKIGGTWGWYAYDPY 284
 QY 201 NLVHGSTAVGPASETGRTGCGTILXGNTFRFARPRDTGIVRHQITPRDNNDQECTE 260
 DB 285 DLFYSSGNPAPWNETMR--PEDNKM-TMAITGREATTGEAKFKAYOKTPHDEWD----- 335
 QY 261 MMTLVNDVOSTMEGLIOSINPAATGERRVLTGVCKTGMWQPDATGEGFLMARDPNY 320
 DB 336 --YAGVAVMLSTBOEDKO-----GQWRKILTHPRDNGIVLYLDTNGSLISA----- 380
 QY 321 QNMIESIDE--NGIVYNEDALIKELDEY-----DVCPTFLGDRPMSALNPD 369
 DB 381 -----DKMDTVNWKVEQDQTGLRPDRPREGTRMDHAKARDICPSAMGHNCHSDYPER 436
 QY 370 GIVFIPLNVCYDMAVDOETSMDYNTSNVTKLRPGKDM-----IGRIDALISTGRT 424
 DB 437 KVFMLGINHICDMERPLLYRAGOFFVAGTILMYGPRATLERAAGQIKRYDAISGM 496
 QY 425 LMSVRAAANTSVLSTGGVLRNGGDRRFRALSOETGETLMQRTLAIVASGOAISTEV 484
 DB 497 KNEKKERSVWGTMATAGGLTFYVTLDTGFKARDSDTGLMLKFKLPGSVIGHPMFYKH 556
 QY 485 DGMQVVAIAGGVSYSG 502

DB 557 DGMQVVAI-----MYGVG 569

 RESULT 9
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC 0924J77
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schober M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-20202376; PubMed-10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT Pseudomonas aeruginosa: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL: AJ009858; CAA0896.1;
DR EMBL: AE004624; AAC05370.1;
DR EMBL: AF068264; AAC79657.1;
DR PDB: 1FLG; 30-AUG-00.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
KW Oxidoreductase; PQQ; Periplasmic; signal; Calcium; 3D-structure;
KW Complete proteome.
FT SIGNAL 1 34
FT CHAIN 35 623
FT DISULFID 139 140
SQ SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;

Query Match
Best Local Similarity 23.2%; Pred. No. 9.2e-24;
Matches 139; Conservative 94; Mismatches 248; Indels 119; Gaps 19;

OY 3 PSLMASAGALAL--LAAPAPQVPTDELLAN--PRAGMISYGOQENYRHSPLT 57
Db 9 PAGLRPSLHCLAEVALAGSAGALANDKDTTGYLQKMGTHAQRWSPK 68
OY 58 QITTEWGOLOLVWARGM---OPKYOVTPLHDGMYTLANGDVIQAIIDAKTGDLWE 113
Db 69 QVNADNVFKLPAMSYSGDEKQKQ--ESQAIYSDGIVYVTSYSLFALDAKTGRLMT 127
OY 114 HRRLP.N-----IATLNSFGIVAN--GVV-----AGST 140
Db 128 YNHRPLPDIPRCOVNNGAIVGDKYFETLIDASYALNKNMGVKKKRAHDGAGT 187
OY 141 QOYSP-----FGCF--VSGHDSATGELMKNYFIR-----169
Db 188 MTGAPTIYKDKTKGVLLIHSGSDEFVGVGRFAPADPTGEGEIMRFEGHMGRLNGK 247
OY 170 -----AGEGEETWGNDEA-----RWMTGA--WGQITVDPVTLVYGSTAAGPASET 216
Db 248 DSYITGVKAPSWPDDRNSTPTGVESWSHGGAPWQSAFPAENITIIYVGNPQWNTW 307
OY 217 QRGTPG-----TLGTNTFAVRPDGELVYRQTLPRNNMOECTFEMATTNVDPOR 270
Db 308 ARAKAGNPDDSLT-TSGQYVDPSSGEVKEWFOHTPNDAMPFOSNNELVLEFYKAD 366
OY 271 STEMEGLQGINR-----AATGERVLTGVPCKTGMW--OFDAETGEFLMARDIYQNM 323
Db 367 GKIVKATAHADRNGEFFVYVDRSGKLONAFPFVNDITWASHIDKTS-----RVEREGQ 421
OY 324 IESIDENGIVYVNEDAIILKEIDVEDYDVPFELGDRDPASALNDGSIYFIPLNVCYDM 383
Db 422 RPLPEPG-----QKHGKAVEVSPPLGGKNNPMAYSDDTGLFYVPANHWKEDY 471
OY 384 MAVDQETSMQVYNTSVNTKLPKGMIGRIDAIDISTGRTLSVRAAANYSPULSTGG 443
Db 472 WTEEVSYTKGSAVLLGMGRFKRMVDHVGSLRAMDPVSGKVVWEHKEHLPLWAGVLATAG 531
OY 444 GVLNNGSTDRFYRALSOETGELTMQTRLATVASGOAISTEVDGMOYVALAGGVSGSL 503
Db 532 NLVFTGTGGDGYFKAFDAKSKELMKFQTGSGIVSPITWEDGDEQYLGVT--VGYGAV 588

RESULT 10
DHML_METME STANDARD; PRT; 571 AA.
AC P38539;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha
subunit) (MDH).
OS Methylophilus methylotrophus (Bacterium W3A1).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;

```

```

OC Methylophilus.
OX NCBI_Taxid=17;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=9405969; PubMed=8241148;
RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
RA Davidson V.L.;
RT "The active site structure of the calcium-containing quinoprotein
methanol dehydrogenase."
RL Biochemistry 32:12955-12958(1993).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=93054513; PubMed=1331050;
RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
RA Mathews F.S.;
RT "The three-dimensional structures of methanol dehydrogenase from two
methylophilic bacteria at 2.6-A resolution."
RL J. Biol. Chem. 267:22289-22297(1992).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +
reduced acceptor.
CC -1- COFACTOR: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
PER TETRAMER.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
DR PDB; 4AAH; 08-DEC-96.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
KW Oxidoreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
KW Calcium.
FT DISULFID 103 104
FT DISULFID 379 408
FT ACT SITE 297 297
FT HELIX 2 9
FT TURN 11 12
FT STRAND 14 14
FT TURN 17 18
FT TURN 21 22
FT STRAND 26 27
FT TURN 34 36
FT HELIX 37 39
FT STRAND 41 47
FT STRAND 59 61
FT TURN 62 63
FT STRAND 64 68
FT TURN 71 73
FT STRAND 75 79
FT TURN 80 81
FT TURN 83 84
FT STRAND 86 90
FT HELIX 96 101
FT STRAND 103 104
FT STRAND 112 114
FT TURN 115 116
FT STRAND 117 121
FT TURN 123 124
FT STRAND 126 131
FT TURN 132 134
FT STRAND 137 142
FT HELIX 146 148
FT TURN 149 149
FT STRAND 151 151
FT STRAND 157 159
FT TURN 160 161
FT STRAND 162 166
FT STRAND 168 168
FT HELIX 170 172
FT TURN 173 173

```


[illegible]

1

DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT_SITE 470 470
 FT VARIANT 788 788
 SO SEQUENCE 808 AA; 87567 MW; 0F4160DA7652445 CRC64;

Query Match Best Local Similarity 10.9%; Score 295; DB 1; Length 808;
 Matches 141; Conservative 83; Mismatches 230; Indels 248; Gaps 25;

QY 13 ALALLAARPAQVPTVDLLA-----NPPAGMISVGGQNGEYRHSPLTQTT 61
 DB 128 AVLAFLASLFTDPDHSIGELPQIANASPADPNVASEMAYIGRTQAGDRWSPINQINA 187
 QY 62 ENYGOLQVW-----AKGMPGKV--OYPLIHGVMYLANPDVIOAIDAKTGLIM 112
 DB 188 TNSNKLKVAWHITKDMNSNDPEQETNATEPIEFNNLTCMLHQKLEAVDAGATGNV 247
 QY 113 EHRRL--PNIATLNSFI-----YANGVIV----- 136
 DB 248 VYDPKQIINPGFQHLRCGVSFHETPRANAMDSGNAPAPDCAKDSILPVDNGRLVEDAD 307
 QY 137 AGSTC-----QYSPGCFV----- 150
 DB 308 TGTCTGCGEENGEIDLRVNPQYPTPGQYEPSTVYIDKLINSAITDNGSVKQASGA 367
 QY 151 -SGHDSATGEELW---RNYFIPRAGEGDEFTGNDYDARMTGAKQOITDPTNLVHY 205
 DB 368 TQAFDYTGKRVWFDASNPQNPJDESHVFNHNSNSIVS---STDANLNLVYI 422
 QY 206 GSTAVGPASEYQGRTP---GGTLYGTNTR---AVRPDGEIYVHQTLPKRDMDQEC 257
 DB 423 PMGV-----GTPDQMGDRKDSERFAPGIALNALNDGKLAWFYQYVHNDIMMEL 473
 QY 238 TFEAMVTVNDYQPSYEMEGIQSINPNATGERRVLTGVPCKTGIMQFDETG-ELLMAR 316
 DB 474 PGPSTLYDVTQKDGTLVPAIYA-----PRTGDIFVLDRTGKEIVPAP 517
 QY 317 DT-----NYQMTESIDENGIVTVNEDAILKEIDV-----EYD 349
 DB 518 ETVPVQGAAPGDHSTPTQPMQO---LTLRPKNPLNDSIDNGTIFPDQMECSIFHLRIE 574
 QY 350 -----VCPFLGGRDMPASALNDPSGIYFPLNNVCYDMAADQ----- 388
 DB 575 GPFTPLSLKSLIFPGDLGFEWKGGLAVDQROVAFAANPLISLPSVQLVPRGPNLME 634
 QY 389 -----EFTSMYNTSVNKLPRCKMI-----GRIDAIDISGRLL 425
 DB 635 ENAKGTGEGTGLQHNHYGIPYAVNLHPLDYLPLFGIKMCRTPPGYVAGIDKTNKV 694
 QY 426 WSVERRAANVS-----PV-----LSTGGGVLE-NGGTDYFRALSOETG 463
 DB 695 WQHNNGTLRBSMTGSSILPLPIKIGVPSLGGPLSTAQNLFLTASMDYIRAVNLMTG 754
 QY 464 ETLMQTRLATVASGOAISYEVNDQYVAIAGGVSYGSLNS 505
 DB 755 KVLMDRLPAGAQATPTIYAINKGQYI-----VYTAGGHS 790

RESULT 14
 QUITA_ACICA STANDARD: PRT; 809 AA.

AC Q59086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (MAD(P)-independent quinate dehydrogenase).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_Taxid:471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsemore D.A., Ornstion L.N.;
 RT "The pca-pob supraprotein cluster of Acinetobacter calcoaceticus
 RT contains quita, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsemore D.A., Ornstion L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus.";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC 1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC 1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
 CC dehydroquinate + reduced pyrroloquinoline-quinone.
 CC 1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC 1- CORFACTOR: PQQ.
 CC 1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC 1- INDUCTION: BY PROTOCATECHUATE.
 CC 1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC
 DR EMBL: J05770; AAC37161.1; -
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 KM Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 SO SEQUENCE 809 AA; 88196 MW; 71P67CEBBA62BFCB CRC64;

Query Match Best Local Similarity 9.4%; Score 253; DB 1; Length 809;
 Matches 146; Conservative 68; Mismatches 245; Indels 220; Gaps 24;

QY 3 PSLILMASGALALLAA--PAFAQVPTVDLLANPAGMISVGGQNGEYRHSPLTQTT 61
 DB 148 PHEYVASEEELPLVVDPAKKV-----NMWHYGDAGGGSFVALDDINR 193

Fri May 24 11:27:30 2002

walick-934-128.pcp.rsp

Page 14

OY 455 FRALSOEGERFLMOTRIATVASGOAISY--EVDGMOYVAIAGG 496
Db 727 LRAFDSATGKELMKGRLPVGSGGPIITYVSHKTKQYVVISAG 770

Search completed: May 24, 2002, 10:30:54
Job time: 632 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:17 ; Search time 93.8 Seconds
(without alignments)
931.369 Million cell updates/sec

Title: WALICK-934-128.PCP
Perfect score: 2705
Sequence: 1 MKPSTLWASAGALALIAAP.....GMQYVAIAGGVSYSGSLNS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvivirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500.5	55.5	608	2	093RE9
2	523	19.3	698	2	09KH03
3	489.5	18.1	742	2	053362
4	481	17.8	708	2	046444
5	478.5	17.7	601	2	09EYH8
6	477.5	17.7	601	16	092WY9
7	474	17.5	695	2	09AQ08
8	470	17.4	629	2	09AC04
9	461	17.0	601	2	P71509
10	459	17.0	691	2	09AF95
11	430	15.9	633	2	024759
12	422.5	15.6	623	2	09AGW3
13	417.5	15.4	599	2	09L935
14	408	15.1	695	2	093460
15	396	14.6	573	2	059540
16	302	11.2	790	2	09X255

17	282.5	10.4	785	16	098KF6	098KF6 rhizobium 1
18	277	10.2	786	2	P95466	P95466 pantoea cit
19	273.5	10.1	777	16	091115	091115 pseudomonas
20	258.5	9.6	703	16	092RB3	092RB3 rhizobium m
21	232	8.6	644	2	052551	052551 pseudomonas
22	227.5	8.4	639	2	P77931	P77931 pseudomonas
23	208.5	7.7	470	2	032699	032699 hyphomicrob
24	206	7.6	470	2	030326	030326 acetobacter
25	203.5	7.5	179	2	032697	032697 hyphomicrob
26	200.5	7.4	180	2	032696	032696 hyphomicrob
27	200.5	7.4	180	2	032700	032700 hyphomicrob
28	200.5	7.4	181	2	032615	032615 hyphomicrob
29	199.5	7.3	181	2	032692	032692 hyphomicrob
30	197.5	7.3	181	2	032703	032703 hyphomicrob
31	196.5	7.3	181	2	032621	032621 hyphomicrob
32	195.5	7.2	182	2	032706	032706 hyphomicrob
33	193.5	7.2	180	2	032701	032701 hyphomicrob
34	191.5	7.1	182	2	032612	032612 hyphomicrob
35	191.5	7.1	182	2	032693	032693 hyphomicrob
36	190.5	7.0	172	2	033884	033884 methylosinu
37	186.5	6.9	172	2	033882	033882 methylosinu
38	186.5	6.9	179	2	032707	032707 hyphomicrob
39	185.5	6.9	180	2	032694	032694 hyphomicrob
40	185.5	6.9	180	2	032704	032704 hyphomicrob
41	185	6.8	171	2	033884	033884 methylosinu
42	183.5	6.8	179	2	032702	032702 hyphomicrob
43	182.5	6.7	184	2	032695	032695 hyphomicrob
44	181.5	6.7	185	2	0918K8	0918K8 uncultured
45	176.5	6.5	172	2	093K58	093K58 methanotrip

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT; 608 AA.
AC 093RE9
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudogluconobacter saccharoketogenes.
OC Bacteria; Pseudogluconobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shibata T., Saito Y.,
RT "Alcohol dehydrogenase."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046580; BAB62258.1; -
DR SEQUENCE 608 AA; 65101 MW; 0ACCE97AE11BA570 CRC64;

Query Match 55.5%; Score 1500.5; DB 2; Length 608;
Best Local Similarity 50.3%; Pred. No. 2.8e-101;
Matches 293; Conservative 63; Mismatches 139; Indels 87; Gaps 6;
10 SAGALALIAAPAFQ-----VPTDELNPAGEWISYGOQEN 50
16 STALIASLSPAPAHDAANAAPSKAGASAIENFQPTADLACKNANPILRGVNG 75
51 YRSPPLQITTEVNGGOLVWARGOPGVQYPLIHGVMYLANPGVYQAIADKGD 110
76 WGYSPLOQINKVDGLQVYMSRTMEGSAIYNGVIFLGNTVYQAIADKGTSL 135
111 IWEHRRQLPNITAT-LNSFG----- 128
136 IWEYRKLPSSAKFINSIGAKRSIALGDKYFVSWDNFVVALDAKTKLAWETNRG 195
129 -----IVANGTVAGSFCQYSPFCFVSGHDSATGEELMKNFTIPRAGEGDET 177

123 ---TLN-----SEGI----- 129
 143 CDTVNGAGYNGKVFNGTFFDRLVAADATGKWEVNTIPADSLGKQRYTVDGAVR 202
 130 VANGVAVAGTCQYSPFGC--FVSGHOSATGEELMRYEIPRAGEBD----- 175
 203 VAKGLVILIN--GSGEFGARGFVSAFDEIETCKLMKRYTVPNNKNEPDHADVNLMSKA 260
 176 -ETMGNDYEARMWT-----GANGQITDPYTNLYHGSTAVGASSETORGTGCTLY-- 226
 261 YKTMGP--KAMVYRQGGGTWDSLVYDVSLLY--LAAG-----NSPNNYKASE 309
 227 -GTN-----TRAVRPDTGEIYWRHDTLPDNDMDQECTFEKMTNVNDVOPSTMEGLQSI 280
 310 GIGSNLGLSIVALKRPTGEIYVHHQATPMDQMDYTSTVOQIMTLDMDV----- 357
 281 NPNMATGE--RRVLTGVPCKTGMQDAETGEFLMARDTYONMIESIDE--NGIVTNEED 338
 338 -----NGEMRVIWHAH--KNGFFYVLDARTGEFLAGKNTVYONNANGIDPLTGRPIYNP 411
 339 ALIKELDEYVCTPFLGGRDMPGSAALNPDGIFYFLNNVCY-----DIMA 385
 412 GYTLTGKFWGIGIPEPLGAGNFKGMATSPKTHLYLPAHQIFGYKNOVGGRKPHDAMN 471
 386 VDOEFTSMQVYNTSNVTKLPKGMIRIDALIDISTGRITLMSVERAANYSPVLSTGGGV 445
 472 VGLDMTKKNGLPDTPPE--ARTAYIKDLHGMILAMDPYKMEIYWKIDHGPWNGSLATGGDL 530
 446 LEPNGSDRFRFALSOETGLMTQRLATVASGAISTEVDGMOYVA-----IAGG 495
 531 LEPGLANGEFHAYDANSGDLYKFPDAGSLIASPMTYSVNGKOYVAEVGWSGITPISMG 590
 496 GVSYSYG 502
 591 GVGRTSG 597

RESULT 4
 046444 PRELIMINARY: PRT: 708 AA.
 ID 046444
 AC 046444
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE QUINOHEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 (EC 1.1.99.-) (OH-BDH1).
 GN QHEBH.
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
 OX NCBI_TaxID=285;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15667;
 RC MEDLINE=96184549; PubMed=8654419;
 RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 De Vries S., Duine J.A.;
 RA "Characterization of the gene encoding quinoxaloproduct ethanol
 dehydrogenase of Comamonas testosteroni";
 RT Eur. J. Biochem. 235:690-698(1996).
 RL [2]
 RN SEQUENCE OF 32-54 AND 477-490.
 RP STRAIN=ATCC 15667;
 RP MEDLINE=95324580; PubMed=7601151;
 RA De Jong G.A.H., Geerloff A., Stoorvogel J., Jongejan J.A., De Vries S.,
 Duine J.A.;
 RA "Quinoxaloproduct ethanol dehydrogenase from Comamonas testosteroni.
 Purification, characterization, and reconstitution of the apoenzyme
 with pyridoxal phosphate analogues";
 RT Eur. J. Biochem. 230:899-905(1996).
 RL [3]
 RN CHARACTERIZATION.
 RP MEDLINE=86242113; PubMed=3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;

RT "Quinoxaloproduct ethanol dehydrogenase apoenzyme from Pseudomonas
 testosteroni";
 RT Biochem. J. 234:611-615(1986).
 RL [4]
 RN CRYSTALLIZATION.
 RP MEDLINE=21536088; PubMed=11679760;
 RA Oubrie A., Huijzinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 Duine J.A., Dijkstra B.W.;
 RA "Crystallization of quinoxaloproduct ethanol dehydrogenase from
 Comamonas testosteroni: crystals with unique optical properties";
 RT Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POO, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -.
 DR HSP: 0924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR00329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_c_1.
 DR PRINTS: PR00605; CYTOCHROME_C1.
 DR SIGNAL: PQO; Heme; Calcium; Oxidoreductase; Periplasmic.
 KW SIGNAL: 1 31
 FT CHAIN: 32 708
 FT BINDING: 635 635
 FT BINDING: 638 638
 FT METAL: 639 639
 FT SEQUENCE: 708 AA; 76822 MW; 99AB54BDD6ACB3 CRC64;

Query Match 17.8%; Score 481; DB 2; Length 708;
 Best Local Similarity 25.4%; Pred. No. 9, 6e-27;
 Matches 150; Conservative 84; Mismatches 239; Indels 118; Gaps 19;
 3 PTLNLSAGALALLAPAFQVPTV-----DELANPP-AGFMISYGNQEN 50
 12 PGRVWVLLAACLG--SNAFAQTGPAAQAAAVCRVDDFLRAAARTPMTITGYDAE 69
 51 YRHSPLQITTEVNGQLVWARGMPGK--VQVPLIDHGVYLANPGVYQADAKTGD 109
 70 TRYSRLQDQANANVKDGLAMSYNLESTRGVATPVVDGIMYASWSVYHAIDRTGN 129
 110 LIWEHRDLPP-----IA----- 122
 130 RIWYDPOIDSTGFKCCDVNRGVALMKKYYVGAMGDRLLDAATGKWHOMTFE 189
 123 -----TLNFGIYANGVYAGST--COYSPFGCVSGHDSATGEELMRYEIP----- 168
 190 GKGSLITIGAPRFKGVIIIGKGAETGYRG--YIRAYDAETGERKWRFSVGPSPKF 248
 169 ---RAGEGDETWGNDYEARMWTGA-----WGQITDPYTNLYHGSTAVGASSETORGT 221
 249 EDSMKRAARTM--DPSCKWEDAGGGGTMDSMTFDELMTLVGTGNSPSSHVRSKP 306
 222 GGTLYGTNTRFVAVRPDTGEIYWRHDTLPDNDMDQECTFEKMTNVNDVOPSTMEGLQSI 280
 307 GGNLYLASIVLADPPTGKYKHHYETPGDNMDYTSTQPMILADIKI----- 353
 282 PNATGRRVLTGVPCKTGMQDAETGEFLMARDTYONMIESIDE--IYTVNEEDAI 340
 354 ---AGRKRYLIHAP--KNGFFYVLDRTNGKFLSAKNVYPVWVAGSYKHKGPPIGI---AA 406
 341 LKELDEYVCTPFLGGRDMPGSAALNPDGIFYFLNNVCYDMAVQD--EFT----- 391
 407 ARDGSKPDVAVPGPYGAHNMHPSFNPOTGLYLAQNVVAVNLMDKKWEPNDAAGPKRQ 466

QY 392 SMVYNTSNVTKL-PPGKMIGRIDALDSTGTMTSVERAANSPVLSTGGVLFNG 450
 Db 467 SGTGMNTRAKFPNPPKSKPPGRLAMDPVAOKAMSVSEHVSPPMNGTLTLAGVYVFGT 526
 QY 451 TDRFRLASOETGETIMOTRLATVASGQAISEYVDGMQYVALA-GGGVSG 500
 Db 527 ADGLVAVYHAHATGEKLEAPGTGVVAAPSTYMDGRQYVAVANGGVG 577

RESULT 5
 Q92YW8 PRELIMINARY: PRT: 601 AA.
 ID Q92YW8
 AC Q92YW8
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-DEC-2001 (TREMBLrel. 16, Last sequence update)
 DE METANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
 GN MKAP
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RM1021;
 RA Fennel B.J., Tiwari R.P., Dilworth M.J.
 RT "Regulation of C1 assimilation in Sinorhizobium meliloti."
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF039488; AAC31643.1;
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR002372; Bac_PDO_repeat.
 DR Pfam: PF01011; Bacterial_PDO; 7.
 SQ SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 17.7%; Score 478.5; DB 2; Length 601;
 Best Local Similarity 25.8%; Pred. No. 1.2e-26;
 Matches 155; Conservative 90; Mismatches 210; Indels 145; Gaps 24;

QY 14 LALLAARAPAOVTPVDEL-LANPAGEMISTGQNGENRHSPLQTITTEWQLOLW 71
 Db 8 LAIMISGGAQVAFANDELQKLDLP-NQMAIQTDVLANLRSKLDQINKDKVGLQVAM 66
 QY 72 --ARGMOPGKQVQVPLIHGVMTLANP-GDVIAQID-AKTGDLMEHR-RQLPNATLNS 126
 Db 67 TFSGVLRGH-EGSPLYVIGDLMVHTPPPTVYALDLSKDGQIVKYPKODPNVIVMC 125
 QY 127 FGIYANGVYVAG-----STCOYSP----- 145
 Db 126 CDVTNRCVAAADNKIFLHQADTTVALDAKTGKIVMSKNGDATKGETNATATVAPYDKI 185
 QY 146 -----FGC--FVSGHDSATGEELMRNYFI-----PRAGEGDET 177
 Db 186 LVGISGGEVGRGHVATAYSMADGKVLKMGSPDSQTLIDPEKTHLGRKPGVDSGLTT 245
 QY 178 WGNDEYARMMTG---AMQIITDPTNLYVHGSTAVGPASETOGRTPGGTLGYNTRFAV 234
 Db 246 WEGD---QMKIGGTTMGWYSYDPEENLYVYGTGNPSTWPTOR--PGDNRM-SMTIFAR 299
 QY 235 RPDTEIYWRHQTLPNDMDQECTFEEMAVTNDVOPSTEMEGLOSINPNAATGERRVLTG 294
 Db 300 DVDITGMAKMLYQMTPHDEMVDYDGVNEMILTGQIDGK-----DRKLILTH 343
 QY 295 VPCKTGTMQFDEATGEFL-----WADRTNYQNMIE-----SIDENGIV 333
 Db 344 FD-RNGFGYTMVDTVELLVAEKDPTVNMATEVMDPDKSGRQVAVQYSTEDNG-- 400
 QY 334 TVNEDAILKEIDVEYDVCPTFLGGRDMPSAALNDPDSGIYFIPLNNVCYDMAVADQETSM 393
 Db 401 ---EDNTNT-----GVCPAALGTCKDQOPAAVSPKTELFYVPTNHVCMQDYEPFRVSYTAG 451
 QY 394 DVTYNTSNVTKLPBGKDM---IGRIDALDSTGTMTSVERAANSPVLSTGGVLFNG 450

Db 452 QPYVGAITLSMYP-KDSHGCMGNTIACDNKESGKIRKSLPPEFVSWSGALATAGDVYFCT 510
 QY 451 TDRFRLASOETGETIMOTRLATVASGQAISEYVDGMQYVALA-GGGVSG 505
 Db 511 LBSYTLAAVDAATGKELVREFKTPSGVYGVNMTYAREGQYVAVLSGVGAGIAGLITN 570

RESULT 6
 Q92WY9 PRELIMINARY: PRT: 601 AA.
 ID Q92WY9
 AC Q92WY9
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
 GN SMO20173.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid PSYMB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Flann T.M., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorheeler F.J., Hernandez-Lucas I., Becker A., Cowie A., Guzy J.,
 RA Golding B., Piehler A.,
 RT "The complete sequence of the 1,683-kb psymb megaplasmid from the N2-
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AL603642; CAC48573.1;
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424FEED13ADB6 CRC64;

Query Match 17.7%; Score 477.5; DB 16; Length 601;
 Best Local Similarity 25.8%; Pred. No. 1.4e-26;
 Matches 155; Conservative 90; Mismatches 210; Indels 145; Gaps 24;

QY 14 LALLAARAPAOVTPVDEL-LANPAGEMISTGQNGENRHSPLQTITTEWQLOLW 71
 Db 8 LAIMISGGAQVAFANDELQKLDLP-NQMAIQTDVLANLRSKLDQINKDKVGLQVAM 66
 QY 72 --ARGMOPGKQVQVPLIHGVMTLANP-GDVIAQID-AKTGDLMEHR-RQLPNATLNS 126
 Db 67 TFSGVLRGH-EGSPLYVIGDLMVHTPPPTVYALDLSKDGQIVKYPKODPNVIVMC 125
 QY 127 FGIYANGVYVAG-----STCOYSP----- 145
 Db 126 CDVTNRCVAAADNKIFLHQADTTVALDAKTGKIVMSKNGDATKGETNATATVAPYDKI 185
 QY 146 -----FGC--FVSGHDSATGEELMRNYFI-----PRAGEGDET 177
 Db 186 LVGISGGEVGRGHVATAYSMADGKVLKMGSPDSQTLIDPEKTHLGRKPGVDSGLTT 245
 QY 178 WGNDEYARMMTG---AMQIITDPTNLYVHGSTAVGPASETOGRTPGGTLGYNTRFAV 234
 Db 246 WEGD---QMKIGGTTMGWYSYDPEENLYVYGTGNPSTWPTOR--PGDNRM-SMTIFAR 299
 QY 235 RPDTEIYWRHQTLPNDMDQECTFEEMAVTNDVOPSTEMEGLOSINPNAATGERRVLTG 294
 Db 300 DVDITGMAKMLYQMTPHDEMVDYDGVNEMILTEQOIDGK-----DRKLILTH 343
 QY 295 VPCKTGTMQFDEATGEFL-----WADRTNYQNMIE-----SIDENGIV 333
 Db 344 FD-RNGFGYTMVDTVELLVAEKDPTVNMATEVMDPDKSGRQVAVQYSTEDNG-- 400
 QY 334 TVNEDAILKEIDVEYDVCPTFLGGRDMPSAALNDPDSGIYFIPLNNVCYDMAVADQETSM 393
 Db 401 ---EDNTNT-----GVCPAALGTCKDQOPAAVSPKTELFYVPTNHVCMQDYEPFRVSYTAG 451

[illegible][illegible]

PRT: 629 AA.

[illegible]

Fri May 24 11:27:31 2002

walick-934-128.pcp.ispt

Query Match
Best Local Similarity 17.0%; Score 461; DB 2; Length 601;
Matches 153; Conservative 102; Mismatches 218; Indels 124; Gaps 26;

Db 1 MKPTSLMASAGALLAFAFAOVPTIDELANPAGEMISTGONENYRHSPLTOT 60
1 MRVHLLALGAGLAA-ASBALANESVLK-VAAP-ABOVLOTVDYANTRISKLDIOIN 54
61 TENGOLOLV-ARGMOPKQVTPILHDGVYLANP-GVIOAIDAKTG-DLIMEHR- 115
55 ASNVKALQVAFWTFSTGVLKRGH-EGSPLVGVNIMVHPFNITVYALDIDQAKIVKPEP 113
116 ROLPNA-----TLNFTGIVANGVY-----AGS 139
114 KODPSVLPVWCCTVNRGLAYADGAILLQADTVLSIDASGVNWSVKNDSKGETN 173
140 TCQYSP-----FG-CFVSGHDSATGSELMRYFT----- 167
174 TATVLPVADKYVIGISGEGEVQCHVTATIDKSGKVMWNGISIGDOLLIVDEKTSIG 233
168 -PRAEEDDETWNDEFAAMMG-AMGITYDDPTNIVHGSTAVGAPASTOGTGG 223
234 KPIAKDSLKTWEGD-QMKTGGGCTKMGWFSYDPKLIDAKTSG--NPSIWNKORPFG 288
224 TLGTATREAVAPPPGEIVNRHDTLPBGNMDDECTFPMVNTVNDVOSTEMEBLOSINP 283
289 NKW-SMTIWMANBDTOMAKWYOMTPHDEWDFDGINEMILTD-----KFPDG----- 334

RESULT 9
P71509 AC P71509 PRELIMINARY; PRT; 601 AA.
ID P71509
AC P71509
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-DEC-2001 (TEMBLrel. 02, Last sequence update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
GN KAP. ANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
OS Methylobacterium extorquens
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
NCBI_taxonomy: 408; Methylobacterium.
RP SEQUENCE FROM N.A.
RX MEDIAN-AM1.
RA Chistoseptova L.
RT Molecular and mutational analysis of a DNA region separating two
RL EMBL: 072662; AAB8890.1; Methylobacterium extorquens AM1.
DR HSP: 072662; AAB8890.1; Methylobacterium extorquens AM1.
DR InterPro: IPR002372; Bac_PPO_repeat.
DR Pfam: PFO1011; Bacterial_PPO_7.
SO SEQUENCE 601 AA; 64952 MW; 68E45C7059CB239 CRC64.

Query Match
Best Local Similarity 17.0%; Score 459; DB 2; Length 691;
Matches 145; Conservative 82; Mismatches 208; Indels 114; Gaps 19;

Db 39 GENISTGONENYRHSPLTOTITENYGOLOLVNARGMOPK-VQVTPILHDGVYLANP 97
41 GEMKTHGDDAGTRISLAQITPPMAKELIWSYDLESSGVAETPIVVDGVMVTAFW 100
98 DYIOAIDAKTGDLIMEHRROLPIATLNS-FGIYANGVY-AGSTCOYSPFGCVSGHDS 155
101 SYVHALDVSSGKRLMTIDPEVPREKGNACCDVNRGVANHECKYFVSLDGRIVA-IDA 159
156 ATGEELM-RNFI-----PRA-GEEDETWGNVDEYBAMMGAM- 193
160 FGRKRWERTLIDDKPYTTGAPRYINGKAVIENGGAEGVAGYITAYDPTASRPV 219
194 I-----TYDPTNIVHGSTAVG-----GO 193
220 VPBGPSLPEDASMAAKTWDPAAGVLSGRRHRYELDLGKYRAGCCISAPAPPS 279
212 PASSTQRTGGLTYNTNRAVBPDDGELVNRHDTLPBGNMDDECTFPMVNTVNDV-Q 289
280 PMSIRKSPAGDNLTYASTAYALRPDTGEYVNHQOFRKADNDYSTODLLADLIEIGK 339
270 PSTEMEGLOSINPMAAGERRVLTGPKTGTMMQDAGTGFELMADTNYOMMESIDE 329
340 P-----RKVILHAP-KNGFEFVIDRTDCKFTSKAQNENYVPMATGDE 380
330 NGIYTVN-EDAILKELDEYDVCTPLGGRDWPSSAALNPSGTYFPILNVCYDMAVDO 388
381 NGRIENPEGAMPGLHMR-----PAPSARTMHSMSYSPTOLAFPAONILP-VLOEDK 436
389 EFTSMYVYNTSN-----VTLPPGKDMIGRIDALIDISTGRTILMSVERAAA 433

RESULT 10
O9AF95 AC O9AF95 PRELIMINARY; PRT; 691 AA.
ID O9AF95
AC O9AF95
DT 01-DEC-2001 (TEMBLrel. 17, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE BDH.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
NCBI_taxonomy: 86174; Pseudomonas.
RP SEQUENCE FROM N.A.
RX Varginal A.S., Alp D.J., Sayavedra-Soto L.A.;
RA characterization of the expression of two distinct alcohol
RT butanogenesis involved in butane metabolism in Pseudomonas
RL submitted (NOV-2001) to the EMBL/GenBank/DBJ databases
DR EMBL: AF355798; AAK47220.2; Pseudomonas
SO SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDMA14B64 CRC64.

OY		265	NVDVOPSTEMEGLSINPNATGERFVLTPGCKTGTMMQFAETGEFLMA---	RDNTNQ	321
	:	:	:	:	:
Db	E-----QIDKE-----		GKKRLTLHPDRNGIVYTLDERNDLLISADKLDT--V	389	
	:	:	:	:	:
OY	322	NMIESIDENGITYVNEDALILKELDVE-YDVCEPTLGGSDWPSAALNPDSGIYEPIPLNNVC		380	
	:	:	:	:	:
Db	390	NVRKHNDLKSGLPVRBDEPFGTRMDRKGETICPSANGHNCDSYDPTKOLFPMGJNHLC		449	
	:	:	:	:	:
OY	381	YDMAVADOETSMDVYNSTNVTKLE-PEGKM----	IORDAIDISGRILMSVERAAAN	434	
	:	:	:	:	:
Db	450	MDEPEMLPRAGQFVGATIMMYPGPKCDRONYLGIQIAKVAALITTYXMEIMERFSV		509	
	:	:	:	:	:
OY	435	YSVPILSTGGVLENGGTDRTFRALSGETEMOTRIATVASGSAISYEWDGMQYALAG		494	
	:	:	:	:	:
Db	510	WGCTLATAGAILVFEGTLDGFLKARNSDTGELLMKRHLPSGVIGYPMTVEHKGVOYIAVMS		569	
	:	:	:	:	:
OY	495	G-GVSYGSGL	503		
	:	:	:	:	:
Db	570	GVGMGPVGL	579		
	:	:	:	:	:
RESULT	12				
O9AGW3	PRELIMINARY;	PRT;	623 AA.		
ID	O9AGW3				
AC	O9AGW3:				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	1-BUTANOL DEHYDROGENASE BOH PRECURSOR.				
OS	Pseudomonas butanovora.				
Bacteria;	Proteobacteria; gamma subdivision; pseudomonadaceae;				
OC	Pseudomonas.				
NCBI_Taxid=86174;					
RN	[1]				
RA	SEQUENCE FROM N.A.				
RP	Sayavedra-Soto L.A., Vangnai A.S., App D.J.;				
RT	"Characterization of the expression of two distinct alcohol				
RT	dehydrogenases involved in butane metabolism in Pseudomonas				
RT	butanovora."				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AF346086; AKK15506.1; -.				
DR	HSSP; Q9Z4J7; JPLG.				
DR	InterPro; IPR002372; Bac_PQQ_repeat.				
DR	Pfam; PF01011; Bacterial_PQQ_4.				
KW	Signal.				
FT	1 29 POTENTIAL.				
SEQUENCE	623 AA; 67553 MW; A573AlC64Ac12D55 CRC64;				
Query Match	15.6%; Score 422.5; DB 2; Length 623;				
Best Local Similarity	24.8%; Pred. No. 1.5e-22;				
Matches 157; Conservative	80; Mismatches 220; Indels 175; Gaps				
OY	2	KPTSLMASGALALLAAPAFACQVTPYDELIAN--PPAGWMISYGONOENYRHSPLOTI	59		
	:	:	:	:	:
Db	7	KPALRLRIYATTAALSLPA-AAAYTDVWEDIANDKHTTGTVLVYLGILKQRHSPLKAI	65		
	:	:	:	:	:
OY	60	TTEVNGOQLVMA-----RCMGKQVQVTPLIHGDVMYLIANGVIOADIKDGDLI	111		
	:	:	:	:	:
Db	66	NTDVAVNALVPAMSFSGGEGEKORQE---AQV--LVHDGVYATASYSKIFALDARSGRKL	120		
	:	:	:	:	:
OY	112	WEHRROLPNATIATLNSFCIVANGVIYAGVICQVSPECFCVSHDSATGELMNRF----	166		
	:	:	:	:	:
Db	121	WEYNARLPD-DIRPCCVNRGAIIYGDKYFFETLDAANVALDRKTGKVYMRKRFGDHKV	179		
	:	:	:	:	:
OY	167	-----IPRAGEGDDTY--	178		
	:	:	:	:	:
Db	180	GYYTMGAFPYIKDOKSQTILLVHGSSGDEFVVGWLFARPDPDGGEEVYMARPVWEGHMRL	239		
	:	:	:	:	:
OY	179	-GND-----YEARMTGA--WCQIITYDPVTNLVHGSTAVAP	212		
	:	:	:	:	:
Db	240	NGKDSYPTGDPKPAPSWPDDPNSPTGKVEA-WISGGGAPMOTASFDEVENNANNVIGAGNPAP	298		
	:	:	:	:	:

```

QY 213 ASETORGTGPG-----GLTGTNTFAVAPDPTGELTWMHQTLPDNDMDOCTEEMVTN- 265
D 299 WNTWKFAPGDDPRNDMSLE-TSGAAYVDASTGELKGYOHPDAMDPSGNNSSVLEEX 357
QY 266 -----VVOPTSTEMEGLQSIIN-----PNAAGGERVLTGCPCKTGTMMO-- 304
D 358 KDRKGTGMVNASAHADRNFPPVTRDMLAKAGCYPKKPTIS--LGAMPEFVDGITWASG 414
QY 305 FDAETGEFLMARDYVQNNMIESIDENGIVYNNEDAILKELDEVDYDCTPFLGGRDPSAA 364
D 415 FDLTKTKPI-EKDNRPQPKGADKGESIFVS-----PEFLGSTMHMS 458
QY 365 LNDGSGIPIPIANVCYDMMAVDOETSMVNTSNVT-----KLPG-----KMT 411
D 459 YSPDTGLFYIPANHWAMD-----YTWENTYTKKASATLGGCFRIKMLFEDHV 505
QY 412 GIDAIDISTGRTLMSVERAANSPVLTSGGVLFNGSTDRYFRALSQETGETLMOTRL 471
D 506 GILRAIDPSPARSLAGAGRVPA-VAGTILTTAGGWFTGTSIDGLKAFDAKNGKELMKFQT 564
QY 472 ATVASGOAISYEVDGMQYVYALAGGVSYSGSL 503
D 565 GSGVYVPTWEMDGEQYVALIOG---YGGAV 593

RESULT 13
Q9L935 PRELIMINARY: PRT: 599 AA.
AC Q9L935;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
GN MXAF.
OS Methylovorus sp. (strain SSI / DSM 11726).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylovorus.
OX NCBI_TaxID=81683;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SSI;
RC Kim Y.M.;
RT "Cloning and nucleotide sequence of mxaf gene of Methylovorus sp.
RT strain SSI DSM11726."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF184915; AAD56237.2;
DR HSSP; P38539; 4AAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO; 7.
DR PROSITE: PS00364; Bacterial_POO_2; 1.
SQ SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

```

```

Query Match 15.4%; Score 417.5; DB 2; Length 599;
Best Local Similarity 24.4%; Pred. No. 3.3e-22;
Matches 142; Conservative 97; Mismatches 229; Indels 115; Gaps 21;

```

```

QY 4 TSLWASGALALAAAPFAQVTPVTDDELANPAGEMISTYQONENRHSFLQITTEN 63
D 6 TATGAAVAGLALSAMLPSSVAADSLAALGANP--NNMTMQIGDVTGGHYSRLSQTITGN 63
QY 64 VGOLOL---WVARMGOPKQVOYVELLHDGVMTLAN--PDVIOAIDATGGLLWEHR--RQ 117
D 64 VKNLESGIGHFSTGLTGH--EGAPLVIGDMYINTPFNNFALNALAPEKIVYQHKPKO 122
QY 118 LPNATLANSIGYANGV-----IVA-----GSTCOY 143
D 123 DASKAAVACDDIYNRGLAYDGHIFKTLQLDHIVASDAKTKGKELMKMNCDDPAVGSITIQ 182
QY 144 SPF-----GC-----FVSGHDSATGEELMKNYFL-----PRA 170
D 144 SPF-----GC-----FVSGHDSATGEELMKNYFL-----PRA 170

```

```

D 183 APFAVAKVLVCGSGAELGVRGYATAYDQKTGELVMSFATGPDNELNAKOPNKDNPHY 242
QY 171 GEE--GDETWGNDYBARMTGA---WGOITPDPTNLVHGSTANGAPASGTGRTGTL 225
D 243 GOMGLGTKTWEGE---GWRKIGGTNKGWYADYDKRLMLFYGGNAPVNETHR--PGDNK 297
QY 226 YGNTNFAVRPDTGELIYVNRHQTLPDNDMDOCTEEMVTVNVOPTSTEMEGLQSIINPAA 285
D 298 W-TWTIMARDVDITGAAMKGYOKTTPHDEMDPAGVQNMILTDQAVNCKTO----- 344
QY 286 TGERVVLTCVPCKTGTMMQDPAETGEFLMARDTN--YQNNMIESIDENGIVYNNEDAILKEL 344
D 345 -----PLTHVD--RNGIYLTNROTGSIVQAKVDAVAVFKVDKTGLPVRDPEESTRM 399
QY 345 DVE-YDVCEPFLGGRDMPSSAALNPDGIFITPLANVCYDMMAVDOETSMVNTSNVTK 403
D 400 DHRGTNCPSAMGFHNQIGDAIDPDSRTTFYGLNHCIDMEPEMLPYRAGQFFVGATLAM 459
QY 404 LP---PGKMTIGRIDAIDISTGRTLMSVERAANSPVLTSGGVLFNGSTDRYFRALS 459
D 460 YPGNPGPTKEMGOVYALMDGVYGEVKKWTKMEKFSWNGTILATGGLVFTNTLDGNIKALD 519
QY 460 QETGETLMQTRLATYASGOAISYEVDGMQYVYALAGGVSYSGS 502
D 520 KTKGKEIMKFKMPSGALIGAPMSYAVKQYIA-----TNYGVG 557

```

```

RESULT 14
Q934G0 PRELIMINARY: PRT: 695 AA.
AC Q934G0;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE LUPANINE HYDROXYLASE PRECURSOR.
GN LUDH.
OS Pseudomonas sp. DH2001.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=163560;
RN [1]
RP SEQUENCE FROM N.A.
RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
RA Rogozinski J.;
RT "Cloning, sequencing and analysis of the gene for Lupanine
RT hydroxylase, a quinoxalochrome c from a Pseudomonas sp."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ318095; CAC67410.1;
DR Signal.
FT SIGNAL.
FT CHAIN 1 26
FT SEQUENCE 695 AA; 74956 MW; 9096C6387E457FEO CRC64;

```

```

Query Match 15.1%; Score 408; DB 2; Length 695;
Best Local Similarity 25.8%; Pred. No. 2e-21;
Matches 139; Conservative 71; Mismatches 218; Indels 110; Gaps 19;

```

```

QY 38 AGEWISTGQNOENYRHSPLQITTEVNGQLDYMARQOPKQVOY--TPLIHDGVMTLANP 96
D 36 SGWMSLLGGGNEQHYFSALKDVYKSNVKNLGLSWPTDMEAGDGLVGNPLVAQDVIRYGGP 95
QY 97 GDVIOAIDATGGLLWEHRQLPNATLANSFGI---VANGVIV-----ASTCO----- 142
D 96 PGKIYVANDKTGKNLMTYTYPEVOYDKTSSWTGFGCTHVNGGLAVDDNNYISGYCKLLAV 155
QY 143 -----YSPFGC-----FVS-----GH-----DSATGEEL 161
D 156 SRTTHKLFWSSGQOPKMKQAITGAPVGGKVFIGNASGDFGGRGHDAADARTGKHL 215
QY 162 WRNYFLPARGE-----EGDETWNNDYBARMM-----TGAMGQITDPTVNLVHY 205
D 216 WRFTYMP--GDPSKPFENDLLAKASKTWGTDY---WKYIKGCVSPWDAITTYDEASDTLYP 270

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:20 ; Search time 101.42 Seconds
(without alignments)
560.735 Million cell updates/sec

Title: WALICK-934-135.PEP
Sequence: 1 MKPSTLWASAGALALAAP.....GMOYVAGGVSYSGSLNS 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A-Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2710	98.8	579	19	AAW37873
2	2385	87.0	579	19	AAW37874
3	2328.5	84.9	579	19	AAW37876
4	2291.5	83.6	578	19	AAW37875
5	484	17.7	754	21	AAW35987
6	483.5	17.6	738	13	AAW20192
7	482.5	17.6	738	11	AAW13993
8	482	17.6	742	11	AAW05235
9	315.5	11.5	740	20	AAW95019
10	158.5	5.8	443	22	ABG24430
11	146	5.3	824	18	AAW09614

12	144.5	5.3	1016	22	AAU34349
13	134.5	4.9	715	22	AAW81738
14	134	4.9	910	22	AAW83007
15	133.5	4.9	696	22	AAW82914
16	130.5	4.8	470	22	AAW82701
17	129	4.7	948	22	ABG25909
18	129	4.7	1510	22	ABG21573
19	116.5	4.2	404	22	AAW82772
20	116	4.2	3567	14	AAW4431
21	113	4.1	1042	21	AAW11732
22	113	4.1	1837	21	AAW11726
23	112	4.1	790	22	AAW67230
24	112	4.1	593	11	AAW07014
25	110.5	4.0	480	8	AAW07068
26	110	4.0	649	22	ABW6538
27	108.5	4.0	530	22	AAU12236
28	107.5	3.9	501	18	AAW10645
29	107	3.9	269	11	AAW07006
30	107	3.9	269	15	AAW53295
31	106.5	3.9	501	18	AAW10644
32	106.5	3.9	514	22	AAW95225
33	106.5	3.9	594	12	AAW10005
34	106	3.9	269	10	AAW94788
35	106	3.9	1012	20	AAW16735
36	105.5	3.8	429	22	AAW66542
37	105.5	3.8	429	22	AAW93799
38	105.5	3.8	469	21	AAW56578
39	105	3.8	278	22	AAW82625
40	105	3.8	505	20	AAW16739
41	104.5	3.8	1721	19	AAW48299
42	104	3.8	1429	22	ABW58779
43	103.5	3.8	501	18	AAW26108
44	103.5	3.8	501	18	AAW26109
45	103.5	3.8	593	15	AAW62944

ALIGNMENTS

RESULT 1	AAW37873	standard: Protein: 579 AA.
ID	AAW37873	
AC	AAW37873	
DT	10-AUG-1998	(first entry)
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulononic acid; L-ascorbic; inhibition.	
OS	Gluconobacter oxydans.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..23
FT		/note="signal peptide"
FT	Protein	24..579
FT		/note="mature protein"
PN	EP832974-A2.	
XX		
PD	01-APR-1998.	
XX		
PF	11-SEP-1997;	97EP-0115801.
XX		
PR	19-SEP-1996;	96EP-0115001.
XX		
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
XX		

Staphylococcus aur
S. epidermidis ope
S. epidermidis ope
S. epidermidis ope
S. epidermidis ope
Novel human diagno
Novel human diagno
S. epidermidis ope
erya region polype
Cryptosporidium pa
Cryptosporidium pa
Amino acid sequenc
Sequence of polype
Drosophila melanog
Human PRO4340 poly
HPV6 mutant L1 pro
HPV6 mutant L1 pro
Protein G variant.
IgG-binding Strept
HPV6 mutant L1 pro
Human protein sequ
Streptococcus GX78
Protein G variant.
C. trachomatis IGV
Human Interferon-a
Human protein sequ
Human prostate can
S. epidermidis ope
C. trachomatis L2
Cryptosporidium pa
Drosophila melanog
HPV11 mutant L1 pr
HPV11 mutant L1 pr
Streptococcus Prot

DR WPI: 1998-195228/18.
N-PSDB: AAV29051.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1: Pages 35-37; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant organisms can be used to convert suitable substrates to aldehydes, ketones or carboxylic acids, especially to convert L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be converted to L-ascorbic acid by standard procedures. The derivatives of AADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA:

Query Match 98.8%; Score 2710; DB 19; Length 579;
Best Local Similarity 92.1%; Pred. No. 2.3e-225;
Matches 512; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

1 MKPTSLMASGALALLAAPAFQVPTDELTPANPAGEMISYGNONEHRSPLQIT 60
1 mkptslmasagallalaaafavtprdeltpanpagewisgngenyrlspilqit 60
61 TENVGQLOLVARMGKQKQVTPPLIHGVMTANPGVIOAIDAKGDLWEHRQLPN 120
61 tenvgqlolvargmkqkvtpplihgvmtnanpgvvioidakgdlwehrrqlpn 120
121 IATNSFGEPTRGMA-----I 136
121 iatnsfgeptrgmalylgtnyfvswdnhlvaldalgvtlfdvrggedmvsnsqpi 180
137 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETWCNDYEARMTG 196
137 vangvivaagstcoyspfcgsvghdsatgelmrnyftipragegedetwcnDYEARMTG 196
181 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETWCNDYEARMTG 240
181 vangvivaagstcoyspfcgsvghdsatgelmrnyftipragegedetwcnDYEARMTG 240
197 AAGQITVDPVNTLVHGSTAVGPASETORPGTILGTNTTRAPVPPDGETVWRHQTLP 256
241 aagqitvdpvntlvhgstavgpaseetorgtltgtnttravppdgetvwrhqtlp 300
257 RDNMDOCTFEMMYTNNVDOVSTEMEGLOSINMATGERVULGVCKTGTMOFDAET 316
301 rdnmDOCTFEMMYTNNVDOVSTEMEGLOSINMATGERVULGVCKTGTMOFDAET 316
317 GEFIMARDTNYQNNIESIDENGIIVTVEDALTELDEREDVCPFTLGGDMPSALNDPS 376
361 gefimardtNYQNNIESIDENGIIVTVEDALTELDEREDVCPFTLGGDMPSALNDPS 420
377 GIVFIPLNVCYDMMAVDQETSMQVNTNNTKPLPGKDMIGRIDAIDISTGRTLMSVE 436
421 givfiplnvcydmmaVDQETSMQVNTNNTKPLPGKDMIGRIDAIDISTGRTLMSVE 480
437 RAAANISPVLSSTGGVLFNGTIDRYFRALSOETGELMOTRLATYASQALSYEVGKQY 496
481 raaanispvlsstggvlfngtIDRYFRALSOETGELMOTRLATYASQALSYEVGKQY 540
497 VALAGGVSYSGLNS 512
541 valaggvsysglns 556

RESULT 2
AAW37874
ID AAW37874 standard; Protein: 579 AA.
XX
AC AAW37874;
XX

DT 10-AUG-1998 (first entry)

XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

DE Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;

XX aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.

XX Gluconobacter oxydans.

OS Key Location/Qualifiers

XX Peptide 1..23

FT Protein /note= "signal peptide"

FT Protein /note= "mature protein"

PN EP82974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN IA ROCHE & CO AG F.

PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.

XX N-PSDB: AAV29052.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1: Pages 38-40; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant organisms can be used to convert suitable substrates to aldehydes, ketones or carboxylic acids, especially to convert L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be converted to L-ascorbic acid by standard procedures. The derivatives of AADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed.

Query Match 87.0%; Score 2385; DB 19; Length 579;
Best Local Similarity 80.1%; Pred. No. 2.9e-197;
Matches 442; Conservative 32; Mismatches 34; Indels 44; Gaps 1;

1 MKPTSLMASGALALLAAPAFQVPTDELTPANPAGEMISYGNONEHRSPLQIT 60
1 mkptslmasagallalaaafavtprdeltpanpagewisgngenyrlspilqit 60
61 TENVGQLOLVARMGKQKQVTPPLIHGVMTANPGVIOAIDAKGDLWEHRQLPN 120
61 tenvgqlolvargmkqkvtpplihgvmtnanpgvvioidakgdlwehrrqlpn 120
121 IATNSFGEPTRGMA-----I 136
121 iatnsfgeptrgmalylgtnyfvswdnhlvaldalgvtlfdvrggedmvsnsqpi 180
137 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETWCNDYEARMTG 196
181 vangvivaagstcoyspfcgsvghdsatgelmrnyftipragegedetwcnDYEARMTG 240
197 AAGQITVDPVNTLVHGSTAVGPASETORPGTILGTNTTRAPVPPDGETVWRHQTLP 256
241 aagqitvdpvntlvhgstavgpaseetorgtltgtnttravppdgetvwrhqtlp 300

XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 XX 19-SEP-1996; 96EP-0115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N,
 PI WPI: 1998-195228/18.
 DR N-PSDB; AAV29053.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of AADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 XX Sequence 578 AA;

Query Match 83.6%; Score 2291.5; DB 19; Length 578;
 Best Local Similarity 76.9%; Pred. No. 3.4e-189;
 Matches 427; Conservative 37; Mismatches 46; Indels 45; Gaps 2;

1 MKPTSLMASGALALIAAPAFVPTDELAMPAGEWISYQONENRHSPLQIT 60
 1 mktllqssaaallvqlpalaq-tatldemlampgawinyqnyrhspqlqit 59
 61 TENVGOLQVWARGMOPGVYVPLIHGCVWLANPGVQAIADAKTGDLTWEHROLPN 120
 60 adnvgqlvwarqmeagvkiqvlvhdgvmylanpgvdaigdaalqdlwehrrlpln 119
 121 IATLNSFGEPTRGMA-----1 136
 120 latlnsfgeptrgmalygltnvfyvwdnhlvaldstgqvvdvdrgqgtdmvsnsqpl 179
 137 VANGVIVASTGCOYSPFGCFVSGHDSATGEELRNRYFLPRAGEGDETTWGNDEYARMNG 196
 180 vangiavagstgcyfpfgcfvsgghdsatgeelrwnflpragegdetwngndeyarmng 239
 197 AAGQITTPVNLVHGSTAVGPASETRGTPGGLTGTMTARAVRPPDGEYWRKQTP 256
 240 vvgqittpvnlvhygstavgpasevtrgtpgglgtmtaravrpdpgeywrkqtp 299
 257 RDNMOEFTFEMMYTNDVOPSTMEGLSINPAATGERRVLTGVCKTGTMOFDAET 316
 300 rdnmoeftefemmytndvopstmeeglslnpaatgerrvltgvcktgtmofdaet 359
 317 GEELMARTYNQNNIESIDENGIYVVEDALTELDERVDVCPTEFLGGDWMSALNDS 376
 360 geflmartynqnniesidenngiyvvedalteldervdvcpfteflggdwmsalnnds 419
 377 GIYFIPILNVCYDMMAVDOEFTSMVYNTKLPKGMIGRIDADISTGRTLMSVE 436
 420 giyfipilnvcydmmavdofestmvynstkltkpgmigrdadistgtrtlmsve 479
 437 RAAANTSPVLTSGGVLFNGGTRDFRFAISOFTGETLWOTRLAYASQAISTYVDQM 496
 480 ryanaspvltsggvlfnggtrdfrfaisoftgetlwtotrlayasqaistyvdmq 539

OY 497 VAIAGGVSYSGSLN 511
 DB 540 valgrgtsygsnln 554

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 XX
 DF 01-MAR-2001 (first entry)
 XX
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulononic acid.
 XX
 OS Gluconobacter oxydans.
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-1B00736.
 XX
 PR 23-APR-1999; 99WO-1B00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI: 2000-687351/67.
 DR N-PSDB; AAC83153.

Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule isolated from Gluconobacter suboxydans useful for the fermentative production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 Claim 1; Fig 8; 96pp; English.
 CC
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 XX Sequence 754 AA;

Query Match 17.7%; Score 484; DB 21; Length 754;
 Best Local Similarity 25.9%; Pred. No. 8e-33;
 Matches 156; Conservative 81; Mismatches 239; Indels 126; Gaps 16;

OY 14 IALLAAPAFVPTV-----TDELLANPAGENWISYQONENRHSPLQIT 66
 DB 17 lgcacaaatcatspvalaedgtatlnadpdmvsgyrtlysegrtyspldgtikdnasn 76
 OY 67 LOIWARMOGPKVQV-TPLIHGCVWYLANPGDVQAIADAKTGDLTWEHROLPN 124
 DB 77 lklwmyhldtnrggeqtlvvgvmyattnvskmkaldaeagkllwysdpxvpgnadr 136
 OY 125 NSFGEPTRGMAI-----1 136

OY 234 --GTN-----TREFAVRPTDGEIYWRHQTLPDNDMDQECTEFEMVTVNDVOPSTEMEGLOSI 287
 DB 310 gfgsnlfjgsvalkpecegyvwhfgatpmdqytsvqimldmpvk-----EN 337
 OY 288 NNNATGTE--RRLVTVCPKCTGTMOQFDETEGLMARDNTYOMNISID-----EN 337
 DB 359 -----gemrhivhap-kngffvyldekgtelsgknyvqmwangldlqtrpmypnd 411
 OY 338 GIATVNEADILKELDEVDVCPTEFLGSRDPSALNPDGSIYFPLNNYCY----- 388
 DB 412 glytlng-----kfwgyl-pgplgahfnemayspkthlvypahqdpfgykngvggf 463
 OY 389 -----DMAAVDEFTSMVYNTSNVTKLPKPKDMIGRIDAIDISGRILMSVERAANYS 444
 DB 464 kphadsnmgldmktnglpdpe-artayikdlhgwllawdpvkmetvklidhkgpnn 522
 OY 445 VLSTGCVLENGSTDRFALSGETELMOTRLATVASGOAISYEDGMOYVA----- 498
 DB 523 llatggdlilfglangedfaydalngsdlkyktdagsgliappmtyvngkgyvavewg 582
 OY 499 -----IAGGVSYSGSG 509
 DB 583 gylpsmggyvgrtsq 597
 RESULT 9
 AAM95019
 ID AAM95019 standard; Protein: 740 AA.
 AC AAM95019;
 XX 21-MAY-1999 (first entry)
 DE Sorbitol dehydrogenase (SDH) protein sequence.
 XX Sorbitol dehydrogenase; SDH; open reading frame; ORF2; L-sorbose;
 KM D-sorbitol; vitamin C; enzyme.
 XX Gluconobacter suboxydans.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /note="signal sequence"
 FT Protein 25..740
 FT /note="mature protein"
 XX EP897984-A2.
 PM 24-FEB-1999.
 PD 13-AUG-1998; 98EP-0115231.
 PF 21-AUG-1997; 97EP-0114432.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;
 PI WPI; 1999-134646/12.
 DR N-PSDB; AAX21501.
 XX New D-sorbitol dehydrogenase gene and recombinant protein - useful
 PT for production of L-sorbose, an intermediate in vitamin C production
 XX Claim 1; Fig 3A-D; 39pp; English.
 PS
 CC This represents a sorbitol dehydrogenase (SDH) protein. The DNA
 CC encoding the SDH enzyme also encodes an open reading frame (ORF2)
 CC product upstream of the SDH open reading frame, needed for SDH activity
 CC in vivo. Host cells transformed by a vector comprising the SDH DNA
 CC sequence are used for the recombinant expression of the sorbitol
 CC dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 CC producing L-sorbose from D-sorbitol. L-sorbose is an important

CC Intermediate in vitamin C production.
 XX Sequence 740 AA:
 SO
 Query Match 11.5%; Score 315.5; DB 20; Length 740;
 Best Local Similarity 21.4%; Pred. No. 2.8e-18;
 Matches 149; Conservative 79; Mismatches 188; Indels 281; Gaps 33;
 OY 23 AOVPTVDELALNPPAGEMTSYCONENRHSPLTQITTEWYGOLOL--VWAGM--QPG 78
 DB 86 sqvpmagpqsanparqdwaygrddhqrlyspelsipenasklkvafvhtgysyprp 145
 OY 79 KV-----QVPL-IHDSVMTANPGDVIOAIDAKTGDILME----- 113
 DB 146 qvkwaaetlupkvgdgllycsamndlik-lpafgkqlwrrnvdkyslpytaackgv 204
 OY 114 -----HRQL-----FN-----IATLSFGPTGGM 134
 DB 205 tyfssvpepqrchmrllegldmrlhavdaetgdcfnfnggqvmimglgesvpgf 264
 OY 135 AI-----VANGVIAGSTC-----QSPFGCVSGHDSATGELMRNYFIPRAGEGDE 183
 DB 265 vsmclpppyingvvvnhbevldggrwapsq-vlrydaesgkfw----- 309
 OY 184 TWGNDYEARMWTGANGQITPDYPTNLVHSGTAVGPASETQRCPTGCTLYGINTR----- 238
 DB 310 awdvnnsgf-----sqpay-rvltaverrlpqlpdrirgq-----gsrlrpdms 355
 OY 239 -----FAVRPD-----TGEIYWRHQTLPDNDMDQECTEFEMVTVNDVOPS 278
 DB 356 aadyssalrsdaenkvsavvaldvktsprwvltghkdwdydlsqatl----- 407
 OY 279 TEMEGLOSINPNATGERVILFGVCKTGMQFDETEGLMARDNTYOMNISIDEN- 337
 DB 408 mdmpg-----pdqvlvpaalm--pkrgqtfvlidrtcgk-----pilpveerp 448
 OY 338 ----GIY-----TVNEDAI-----LKELD----- 352
 DB 449 apspgvilpgdprspqgsvmpalirvpdketdmwmspidqicrlikfranygeft 508
 OY 353 -----VEYDVCPTFLGSRDPSALNPDGSIYFPLN-NVCTDMANAVDEFISM--- 400
 DB 509 ppsvdkpwley--pynggsdwsmsydpqsgllaanvnlipmydqvltrkadsiglm 565
 OY 401 -----DYNTSNVTKLPKPKDMIGRIDAIDIS 427
 DB 566 plddpntfkpgggagengamdqtrpygivrpfwdqy-tgmncnrrp-----ymitaicmk 620
 OY 428 TG-RTLWS-----VERAANYSEVLSTGGCVLENG-GTDRYFRALS 466
 DB 621 hqkvltwqhpilgtarangepwlpiglpwelsigtpnngsvvlg9gliffgaatdnqitraid 680
 OY 467 QETGETLMQTRLATVASGOAISYEDGMOYVAIAGGG 503
 DB 681 ehgkvvasavlp99gqanpmtlyeanginqyvalmagg 717
 RESULT 10
 ABG24430
 ID ABG24430 standard; Protein: 443 AA.
 AC ABG24430;
 XX 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #24421.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX

PN	W0200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	
XX	Drimanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS88617.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
XX	Claim 20; SEQ ID NO 54789; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, and polymerase chain reaction (PCR) primers, oligomers and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. AB600010-AB630377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Query Match	Best local similarity	Score	DB	Length
Matches	47; Conservative	22; Mismatches	49; Indels	29; Gaps
QY	6	LIMAS-----	ACALALAPAFQVTVYTEL	LANPAGEWISYCONCENTRHSPTLQ
		168	ltwagfndpgeinglttsadctpeaa	ispvpaq-----dwpaygrngdqgrfspi
QY	59	ITTFNQGLQLVNA-----	KQMGRKV--GVTP	LIDGVNYLANGVDAIQALDKATGD
		219	ineadnmllkeawfittgvdvkgpndg	geitnevpikgtdllyctahgrflaldasqk
QY	110	LIMEHRRQLINIALTNSFCEPT--	ROMA	135
		Db	279	ekwtyhd---pelktneisfghvltcivgs
				302

RESULT	11
AAW09614	
ID	AAW09614 standard; Protein; 824 AA
XX	
AC	
XX	AAW09614;
DT	20-MAY-1997 (first entry)
XX	

DE Pathogenic *Staphylococcus aureus* ORF1 gene product.
 XX
 XX Pathogenicity; vaccine; immunoassay; antibody.
 XX
 OS *Staphylococcus aureus* strain NCTC 8325-4 (ISP8).
 XX
 PN W06641878-A1.
 XX
 PD 27-DEC-1996.
 XX
 PE 13-JUN-1996; 96WO-AU00353.
 XX
 PR 13-JUN-1995; 95AU-0003507.
 XX
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Adham El-Adhami W, Matthei KI, Stewart PR;
 XX
 DR WPI; 1997-065465/06.
 XX
 DR N-PSDB; AAT47757.
 XX
 PT Nucleic acids specific to pathogenic *Staphylococcus aureus* strains
 XX for diagnosis and immunogenic protein prodn.
 XX
 PS Claim 30; Page 62-65; 88pp; English.
 XX

PT Nucleic acids specific to pathogenic *Staphylococcus aureus* strains
 PR for diagnosis and immunogenic protein prodn.
 XX
 PS
 XX
 Claim 30, Page 62-65; 86pp; English.

XX An 824-amino acid protein (AAW09614) and 192-amino acid protein (AAW09615) were identified as products of 2 open reading frames in a nucleic acid (AAI47757) isolated from pathogenic *Staphylococcus aureus* strain 1S78. The 2 gene products are highly hydrophilic and have high surface probability and antigenic index measurements, suggesting that they are surface expressed. The polypeptides, their fragments, and antibodies raised against them can be used in vaccines and as immunoassay reagents for detecting pathogenic *S. aureus* strains.

XX Sequence 824 AA;

```

Query Match          5 39; Score 146; DB 18; Length 824;
Best Local Similarity 19.4%; Pred. No. 0.0015;
Matches    90; Conservative   62; Mismatches 125; Indels 186; Gaps 21
OY      19  ARAFAAYT-----PYTDELLANPAPGEMISYGCNOENYRHSPLQITENVGOQLY 70
           |||::|||               ||| ||| |||::|||         :|||
Db      427 apftevntregqkgektlptlpklkplgtislsgskeeltkdpinel----- 476
           |||::|||               ||| ||| |||::|||         :|||
OY      71  WARGMOPEKVOTPTLIHD-----GWVYLANGDVIQAIDAKTCDLIMEHR 115
           :||::|||               :|||::|||               :|||::|||
Db      477 -----eygpeltpghrdefdrpkltprgkeevpyrgyiknp-----etldvv--- 518
           |::|||               |::|||               |::|||
OY      116 ROLPMIATILNFEFGEPTRGMAIVANGYIVAGSTCOYSPPCGFSCHDSANTSDELIMRYFYIP 175
           ||::|||               |||::|||               |||::|||
Db      519 ---rrpvdsvqgyr pyvgkdslvekeelpfkkkerkinp-----dlapglek-----vt 562
           |||::|||               |||::|||               |||::|||
OY      176 RAGEGGDETGMNDYEARMNTG-----AMQOIYDPTLVNHYGSSNAAGA- 220
           |||::|||               |||::|||               |||::|||
Db      563 regdkggektltpklknpiltgelislgskesqeiktkdpinelteygpeltpghrdefdkp 622
           |||::|||               |||::|||               |||::|||
OY      221 -----SEFGRTGTGGGLNYLTNRFRPRDT-----GELVMRHQTLPDRNNDOECT 265
           |||::|||               |||::|||               |||::|||
Db      623 lptgekeevpygkpg--lknpeitgdvrvrrpvdsvtklygvpykgdslvekeelp--fekerk 677
           |||::|||               |||::|||               |||::|||
OY      266 FEMWTVTPVDROPSTE-----MEGIQSI----- 287
           |||::|||               |||::|||               |||::|||
Db      678 f-----npldapglekvrtregqkgektlptlpklkplgtislsgskeeltkdpinel 722
           |||::|||               |||::|||               |||::|||
OY      288 -----NPNAAGGERRVTLGVPCKTCTGMMQGFDEETGEFFLMADTYQMITE 332
           |||::|||               |||::|||               |||::|||
Db      733 eygpeltpghrdefdrpkltprgkeevpyrgyiknpregdvvtppvd 779
           |||::|||               |||::|||               |||::|||
OY      333 SIDENGIVVNEDALIKELDXYEVCTPFLGGSDMFSAALND 375
           |||::|||               |||::|||               |||::|||

```


Db 780 svlykg--pvkgdsivekeelpfe-----kerkimpd 803

RESULT 12

AA034349 standard: Protein; 1016 AA.

AA034349;

14-FEB-2002 (first entry)

Staphylococcus aureus cellular proliferation protein #625.

Antisense: prokaryotic cellular proliferation protein;

antibiotic; antibacterial; drug design.

Staphylococcus aureus.

WO2001.70955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207727P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI: 2001-611495/70.

N-PSDB; AAS52208.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -

Example 3; Seq ID No 5845; 51pp; English.

The invention relates to antisense inhibitors of genes essential to

prokaryotic cellular proliferation, their use in identifying the

genes, their use in the discovery of novel antibiotics, the essential

genes themselves and the encoded proteins. The prokaryotes used are

Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

invention is also useful for the identification of potential new targets

for antibiotic development. The antisense nucleic acids can also be used

to identify proteins used in proliferation, to express these proteins,

and to obtain antibodies capable of binding to the expressed proteins.

The proteins can be used to screen compounds in rational drug discovery

programmes. The antisense nucleic acid sequence is also useful to screen

for homologous nucleic acids which are required for cell proliferation in

essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic

format directly from WIPO at

ftp.wipo.int/pub/published_pcl_sequences.

Sequence 1016 AA;

Query Match 5.3%; Score 144.5; DB 22; Length 1016;

Best Local Similarity 21.3%; Pred. No. 0.0027;

Matches 115; Conservative 63; Mismatches 170; Indels 191; Gaps 28;

QY 19 AFAFAQVT-----PVTDELLANPPAGEMISYCONQENYRHSPLTITTEWGOLOLV 70

Db 528 apylekvtregqgkgtcttltplcknplgvliskgeskeeltkdpinelt----- 577

QY 71 WAROMPGKVOVTPPLIHD-----GYMVLANQGVDAIDAKTGDLIWEHR 115

Db 578 -----eygpctitpghndefqkplptgekeevpgkplnp-----etgdv----- 619

QY 116 ROLFNATLNSFGEPTRGMAIVANGIVAGSTCOYSPFCFVSGHDSATGELMNNYRFP 175

Db 620 --rppvsvlkyg--pvkgdsivekeelpfekeekfnp-----dlapylek-----vt 663

QY 176 RAGEGDETMGNDYEAARMWTA-----WGQTTDPVNTLVHGSTAVGPA----- 220

Db 664 regqgkgtcttltplcknplgvliskgeskeeltkdpineltleygpctitpghndefqk 723

QY 221 -----SEFQRTPGTGLTGTTRFAVRPDT-----GEIWRHQTLPDNDMDQECT 265

Db 724 lptgekeevpgkpg--lknpetgdvvrpydvsvlkygvpvkgdsivekeelp--lkeek 778

QY 266 FEMWVTNVDVOPSTEMEGLQSIINPNAATGERRVLTGVPCKTGMQFDEAGETLWARDT 325

Db 779 f-----nplapyle-----kviregqgkgtcttltplcknpl-----tgeil----- 815

QY 326 NQNMTESTIDENGIVTVNEDATLKEIDVEYDVCPTFLGGRDWPESAL----- 372

Db 816 ---skgesiee-----itkdpinel-eygpctitpghndefqkplptgekeevpgkpg 865

QY 373 --NPDGIVFIFPLNNVCYDMAVDOFTSMQVYNTKRLPPCK-DMIGRIDAIIDISTG 429

Db 866 lknpetgdvvrp-----pvd-----svlkygvpvkgdsivekeelpfeke 904

QY 430 RTLMSVRAAANSPVLSGTGGVLFNGCTDRYFRALSGTCTTLMQTRLATVAGSOAIS 488

Db 905 f-----kinpdlap-----gtekvtre-gqgkgtcttltplcknplgvlis 944

RESULT 13

AA081738 standard: Protein; 715 AA.

AA081738;

03-SEP-2001 (first entry)

S. epidermidis open reading frame protein sequence SEQ ID NO:570.

Staphylococcus epidermidis SRI strain; infection; diagnosis;

Staphylococcus epidermidis.

Staphylococcus epidermidis.

WO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US30782.

09-NOV-1999; 99US-0164258.

(GLAX) GLAXO GROUP LTD.

Kimmerly WJ;

WPI: 2001-316495/33.

N-PSDB; AAS52586.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

useful for vaccinating against infections, e.g. endocarditis -

Claim 18; Page 189; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

(II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.

CC

CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC no sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 SQ Sequence 715 AA:

Query Match
 Best Local Similarity 4.9%; Score 134.5; DB 22; Length 715;
 Matches 105; Conservative 51; Mismatches 156; Indels 141; Gaps 26;

QY 26 TPVTDELLANPPAGEWISYQONENRHSPLQIT-----TENVOLOLWARQMO- 76
 DB 309 tptt-----knpltekgvgepektekvkqvdeivhyggeelkpgkhddefnabpksqg 364
 QY 77 --PKG-----VOVPLIHGVMYLANPGDVIADAKGDLIMHRQL-PNIA-- 122
 DB 365 dvppkpgvknptdtegvvtpvddvltkygvdgpl-----tsteepfdkkretfdnlpag 420
 QY 123 --TLNSGEPTRGMAIVANGVIVAGSTCOYSPFGCFVSGHDSATGELMWRNYFIPRAGEE 180
 DB 421 tekvvqgep-----gtklttpr-----tknpltekgv-----ge 451
 QY 181 GDETWGNDYFARMWGMQIYDPVTNLVHGSTAVGP-----ASETQROGT 227
 DB 452 gept-----ekvtkqvdeivhyggeelkpgkhddefnabpksqgqtkp 497
 QY 228 PGGLTGTNTRRAVRPDTEGIYWRHQTLPRDNMOECTFEEMVYN--VDVOPSTMEGL 284
 DB 498 pg-----vknpdtegvv-----tppvdd-----vtkygvdgdpitsteel 533
 QY 285 -----QSLNPNATGERRV-LTGP-CKT-GTMMQFDEATGEFLMARDNNGMIESIDE 336
 DB 534 pfdkkretfdnlpdkpgeevtkvkgqgeptklttptknpitgekvgegeptektkp 593
 QY 337 NGITVNDAILKELDVYDYVCPFLGGRWPS--AALNPDSG-IYFIPLNWCV----- 388
 DB 594 --lteygeelkpgkhddefnabpksqgedvppkpgvknptdtegvvtpvddvltkygvd 651
 QY 389 -DNMAVNDG--FTSMQDYNTSNVTKLPPGKDM 418
 DB 652 gdpitsteelpfdkkretfdnlp-----ltpkqkerv 680

RESULT 14

AAG83007
 ID AAG83007 standard; Protein; 910 AA.

XX AAG83007;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3108.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KW vaccination; endocarditis.

OS Staphylococcus epidermidis.

XX
 PN WO200134809-A2.
 XX 17-MAY-2001.
 PD
 XX
 PF 09-NOV-2000; 2000WO-0530782.
 XX
 FR 09-NOV-1999; 99US-0164258.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 PI
 XX
 PT Kimerly MJ;
 DR WPI: 2001-316495/33.
 XX N-PSDB: AAH53857.
 PT
 PS Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 useful for vaccinating against infections, e.g. endocarditis -
 Claim 18; Page 819; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC no sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 SQ Sequence 910 AA:

Query Match
 Best Local Similarity 4.9%; Score 134; DB 22; Length 910;
 Matches 98; Conservative 45; Mismatches 147; Indels 108; Gaps 23;

QY 26 TPVTDELLANPPAGEWISYQONENRHSPLQIT-----ENV-----GOLQVWARQMO- 76
 DB 231 tptt-----knpltekgvgepektekvkqvdeivhyggeelkpgkhddefnabpksqg 286
 QY 77 --PKG-----VOVPLIHGVMYLANPGDVIADAKGDLIMHRQL-PNIA-- 122
 DB 287 dvppkpgvknptdtegvvtpvddvltkygvdgpl-----tsteepfdkkretfdnlpag 342
 QY 123 --TLNSGEPTRGMAIVANGVIVAGSTCOYSPFGCFVSGHDSATGELMWRNYFIPRAGEE 180
 DB 343 tekvvqgep-----gtklttpr-----tknpltekgv-----ge 373
 QY 181 GDETWGNDYFARMWGMQIYDPVTNLVHGSTAVGP-----SETQROGPGTLGTNTR 238
 DB 374 gept-----ekvtkqvdeivhyggeelkpgkhddefnabpksqgedvppk 419
 QY 239 FAVR-PDTEGIYWRHQTLPRDNMOECTFEEMVYN--VDVOPSTMEGL-----QSLNPN 289
 DB 420 pgvknptdtegvv-----tppvdd-----vtkygvdgdpitsteelpfdkkretfdp 465
 QY 290 NAATG-ERRVLTGP-CKT-GTMMQFDEATGEFLMARDNNGMIESIDENGITVVEDA 346
 DB 466 nlapgtekvqkggeptklttptknpitgekvgegeptektkpvdv--lvhygee 523

QY 347 ILKELDEYDVCPTFLGGRDWS--AALNPDGITYETP 382
 DB 524 lrpghkdefdpnapkgsgedvpgkpyvknptdgtgevtpr 561

RESULT 15
 AAG82914
 ID AAG82914 standard; Protein; 696 AA.

AC AAG82914;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KM vaccination; endocarditis.

XX Staphylococcus epidermidis.

XX MO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000MO-US30782.

XX 09-NOV-1999; 9905-0164258.

XX (GLAX) GLAXO GROUP LTD.

XX Kimerly MJ;

XX WPI; 2001-316495/33.

XX N-PSDB; AAH53764.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -

XX Claim 18; Page 763-764; 2188pp; English.

PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

CC Sequence 696 AA;

Query Match 4.9%; Score 133.5; DB 22; Length 696;
 Best Local Similarity 23.6%; Pred. No. 0.014;
 Matches 107; Conservative 56; Mismatches 148; Indels 143; Gaps 27;

QY 26 TPVTDELLANPPAGEISTGQNOENTYRHSPLQTIT---ENV-----GQLQLVARGMO- 76
 DB 103 tptt---knplgtkevgsepteiktqpydeltteygeelkpgkhdefdnapksqge 158
 QY 77 --PGK-----VGVVPLIHGVWYLANPGDVIAIDAKTGDLIMEHRRQL-PNIA-- 122

DB 159 dvpkpgvknptdgtgevtprpvdvltkygpydgdpj-----tsteelpfdckretnplapg 214
 QY 123 --TLNSPEEPTRGMALIVANGVITAGSTCOXSPFCFVSGHDSATGEELMKNYPIPRAGEE 180
 DB 215 tekvvqkgep-----gkctltpr-----tknplgtgekv-----ge 245
 QY 181 GDETWGNDYEARMWTGAMGQITTDPTNLVHYGSTAVGPA--SETQGTGGLXGNTNR 238
 DB 246 gept-----ekltkqpvdeivhyggeelkpgkhdefdnapksqgedvpgk 291
 QY 239 FAVR-PDTGEIYWRHQTLPRDNMDQECTFEAMATN---VDVQSTMEGL-----QSINP 289
 DB 292 pgvknptdgtgevtprpvdvltkygpydgdpj-----tknplgtgekv-----ge 337
 QY 290 NATG-ERRVLTVGP-CKT-GTMOQDAETGEFLWARDJTNOMIESIDENGITYVNEA 346
 DB 338 nlapgtkevvqkgepteiktqpydeltteygeelkpgkhdefdnapksqgedvpgk 395
 QY 347 ILKELDEYDVCPTFLGGRDWSAAL-----NDQSG-IYFIPLNNVCYDMA 392
 DB 396 lrpghkdefdpnapkgsgedvpgkpyvknptdgtgevtprpvdvltkygpk 444
 QY 393 VDEFTSMD-----VYNTSNVTKLPPGKDMI 418
 DB 445 vgnplsteeipfdckrvfnpd-----lkpgeerv 474

Search completed: May 24, 2002, 10:19:22
 Job time: 385 sec

Fri May 24 11:27:31 2002

wallick-934-135.pep.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:13 ; Search time 39.53 seconds
(without alignments)
316.365 Million cell updates/sec

Title: WALICK-934-135.PEP
Perfect score: 2742
Sequence: 1 MKPTSLIMNSAGALALLAAP.....GMQVYAIAGGVSYSGILNS 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	17.7	754	US-09-296-284-4	Sequence 4, Appli
2	483.5	17.6	738	US-07-985-458-3	Sequence 3, Appli
3	475	17.3	720	US-09-296-284-25	Sequence 25, Appli
4	320.5	11.7	739	US-09-136-251-2	Sequence 2, Appli
5	116	4.2	3567	US-07-642-734C-4	Sequence 4, Appli
6	116	4.2	3567	US-08-439-009A-4	Sequence 4, Appli
7	113	4.1	1042	US-08-928-361B-11	Sequence 11, Appli
8	113	4.1	1837	US-08-928-361B-5	Sequence 5, Appli
9	107.5	3.9	1721	US-08-700-651-5	Sequence 10, Appli
10	105	3.8	1012	US-08-219-262B-10	Sequence 10, Appli
11	105	3.8	1012	US-09-031-655-10	Sequence 10, Appli
12	104.5	3.7	635	US-08-931-608A-5	Sequence 5, Appli
13	101.5	3.7	1043	US-08-928-361B-30	Sequence 30, Appli
14	101.5	3.7	1721	US-08-928-361B-6	Sequence 4, Appli
15	98.5	3.6	816	US-07-731-157A-4	Sequence 4, Appli
16	98.5	3.6	816	US-08-229-444B-2	Sequence 4, Appli
17	98.5	3.6	816	US-08-541-780-4	Sequence 4, Appli
18	97	3.5	574	US-09-142-623-13	Sequence 13, Appli
19	95.5	3.5	500	US-08-987-519-2	Sequence 2, Appli
20	95.5	3.5	548	PCT-US93-10541-2	Sequence 19, Appli
21	95	3.5	1012	US-08-931-608A-5	Sequence 2, Appli
22	94.5	3.4	512	US-09-356-818A-2	Sequence 3, Appli
23	94.5	3.4	824	US-09-626-589-3	Sequence 2, Appli
24	94.5	3.4	1612	US-08-169-927-2	Sequence 6, Appli
25	94	3.4	974	US-08-868-786-6	Sequence 2, Appli
26	94	3.4	983	US-09-394-200-2	Sequence 2, Appli
27	93.5	3.4	548	US-08-247-902A-2	Sequence 2, Appli

28	93.5	3.4	966	2	US-08-868-786-2	Sequence 2, Appli
29	93	3.4	380	4	US-08-971-782-4	Sequence 4, Appli
30	93	3.4	380	4	US-09-309-026-4	Sequence 4, Appli
31	93	3.4	439	3	US-08-971-782-2	Sequence 2, Appli
32	93	3.4	459	4	US-09-309-026-2	Sequence 2, Appli
33	93	3.4	659	4	US-09-626-589-2	Sequence 1, Appli
34	93	3.4	717	4	US-09-626-589-1	Sequence 6, Appli
35	92	3.4	532	2	US-08-883-534-3	Sequence 3, Appli
36	92	3.4	606	2	US-09-204-764-3	Sequence 2, Appli
37	92	3.4	606	2	US-08-038-682-2	Sequence 2, Appli
38	91.5	3.3	1536	1	US-08-302-832-2	Sequence 2, Appli
39	91.5	3.3	1536	1	US-08-530-198-2	Sequence 2, Appli
40	91.5	3.3	1536	2	US-08-469-850-2	Sequence 2, Appli
41	91.5	3.3	1536	2	US-08-728-470-2	Sequence 2, Appli
42	91.5	3.3	1536	2	US-08-617-697-2	Sequence 2, Appli
43	91.5	3.3	1536	4	US-08-719-641-2	Sequence 2, Appli
44	91.5	3.3	306	2	US-08-824-707-2	Sequence 2, Appli
45	91	3.3				

ALIGNMENTS

RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Chol, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter suboxydans Sorbitol Dehydrogenase. Genes
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296.284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
; US-09-296-284-4

Query Match	17.7%	Score 484	DB 4	Length 754
Best Local Similarity	25.9%	Pred. No. 2.9e-36		
Matches 156	Conservative	81	Mismatches 239	Indels 126
Gaps				16
14	LALLAFAFAQVTV-----TDELLANPPAGEWISGONENRRHSPLQITTEWVGQ	66		
17	LGCAALAFATSPVALAEDTGATITNADQHPGDMWSYGRYSSEORTSPDQITKDNASN	76		
67	LQLVWANGMPKGVQV-TPLIHGVVYLANPGDVIOALDAIDAKTGLIWEHRRLP-NIATL	124		
77	LKLAMHVDLTNRQDQSTPLVDGVVATTTNMSKKALDAATGKILMSYDPRVGNIAADR	136		
125	NSFGEPTRGMAI-----	136		
137	GCCTVNRGAAYWNGKVFYFTGDRILALDAKTGLMSVYTPKEADIGHORSTYDUGA	196		
137	--VANGVIVASTQYSPFGC--FVSGHSATGEELMRVYFIPRAGEEDD-----	182		
197	PRIAKGVIITN--GGAFFAGRGVTAIDETKMKMRFFYPNDNKRKDAASDVILMS	254		
197	ETWGDYDARMWTC--AMGQITVPTVNTLVHVGSTAVGPASTGTGCTGGTILYGN	236		
237	TREAVRPDTGELVHRQTLPPDNDNDOCTEFEMVNTVNDVQPTSTEMGLQISINPNATGE-	295		
315	SIVAINPDTGKVMHFQETPMDDMDYTSVQOIHALLDMPV-----NGEM	357		

QY	296	RRVLTPGKCTGMMQFDETECEEFILMAADTNYOMNIESIDE-NGIATVNEDAILKELDVE	35
Db	358	RHVLVHP-KNGFEYIIDKTKFISGKPYTEEMWANGIDPVTGRENNPDLATLWLNKP	416
QY	355	YDVCPTFLGRDMPKSAALNPDGSIYFIPLNVCY-----DMAVDQFTSM	401
Db	417	WYGLPGDLGSHNFAANAASPTQKLVYTPAQOVFPVDDPOKGFKAHHDSMNIGIDMUNKIG	476
QY	402	VYNTSN-----VTKLPBGKDMTGRIDAIDISTGRTLMSVERAANAANSPVLSGTGGVILNNG	457
Db	477	LILDNDQHKRAKPAFLFDLDLKGVIAMDPOKQAAFTVYHKKPMWNGGLATATGGVILNGL	536
QY	458	TDRFPAALSOETGETLIMOTRLATVAASGAALSYEVDQMZYAI-----AG-----GGVSYG	507
Db	537	ANEEFAIYATATTGKDLFTFPQASAIIPAPVTYATNGKQIVAVEVGWGIYFELDGVAKI	596
QY	508	SG 509	
Db	597	SG 598	

RESULT 2
US-07-985-

US-07-985-458-3
; Sequence 3, Application US/07985458
; Patent No. 5344777

GENERAL INFORMATION

APPLICANT: Tamaki, Toshimi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Tayama, Kenji;
 APPLICANT: Fukaya, Masahiro;
 APPLICANT: Okumura, Hajime and
 APPLICANT: Kawamura, Yoshiya
 TITLE OF INVENTION: structural Gene of Membrane-Bound
 TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
 TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
 TITLE OF INVENTION: Bacteria
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Frischauf, Holtz, Goodman & Woodward, P.C
STREET: 600 Third Avenue

CITY: New York
STATE: New York

COUNTRY: USA
ZIP: 10016-2

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch
COMPUTER: IBM PC AT

COMPUER: IBM PC compatible (NEC PC-9801 ES),
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form

CURRENT APPLICATION DATA:
APPLICATION NUMBER: TS/07/985 459

FILING DATE: 19921203
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221

FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990

FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: C. J. ...

NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER:

REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213)873-1400

TELEFAX: (212) 370-1622
TELEX: 236268

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 738 amino acids
TYPE: amino acid

TOPOLOGY: unknown
 MOLECULE TYPE: protein

FEATURE:

```

1 NAME/KEY: MATURE PEPTIDE
2 LOCATION: 36 to 738
3 IDENTIFICATION METHOD: N-terminal sequences of the
4 IDENTIFICATION METHOD: purified protein having a molecular weight of about
5 IDENTIFICATION METHOD: 72,000
6 ORIGINAL SOURCE:
7 ORGANISM: Acetobacter altoacetigenes
8 STRAIN: MH-24
9 PUBLICATION INFORMATION:
10 AUTHORS: Tamaki, Toshimi;
11 AUTHORS: Fukaya, Masahiro;
12 AUTHORS: Takemura, Hiroshi;
13 AUTHORS: Tayama, Kenji;
14 AUTHORS: Okumura, Hajime;
15 AUTHORS: Kawamura, Yoshiya;
16 AUTHORS: Nishiyama, Makoto;
17 AUTHORS: Horinouchi, Sueharu and
18 AUTHORS: Beppu, Teruhiko
19 TITLE: Cloning and Sequencing of the Gene Cluster
20 TITLE: Encoding Two Subunits of Membrane-Bound
21 TITLE: Alcohol Dehydrogenase from Acetobacter
22 TITLE: polyoxogenes
23 JOURNAL: Biochimica et Biophysica Acta.
24 VOLUME: 1086
25 PAGES: 282-300
26 DATE: 1991
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

Query Match 17.6%; Score 483.5; DB 1; Length 738;
Best Local Similarity 25.8%; Pred. No. 3.1e-36;
Matches 152; Conservative 76; Mismatches 239; Indels 123; Gaps 14.

QY	10	SAGLILLLAARFAOYTPVTDE-----LLANPAGEIYISGONENYRHSPLQITTT	61
Db	17	TAGTICALISGATWASADDOGATGEAIIHADHDHGNMNTYGRITSPDRSPLQIIR	76
QY	62	ENYGOLOLWARGMOPGRKQY-TPLIHGWYKLANPCDYQALDATGOLIIEHRRLD	11
Db	77	SNVGNLAKLAMIYDLDLTNNGDEGTEPLVDGVNATITNMSMKRAVDATGKLMSYDPVPG	131
QY	120	NIATLNSGEPTRGMAI-----	133
Db	137	NIADKCCCTYVRGAAYNKGAVYFGTFDGRLLMLDKTGLKWSVNTTPPEALCKORSY	196
QY	137	-----VANCYIYAGSTCOYSPGC--EYSGHDSATGEELRNRYFIPRAGEGD-----	187
Db	197	IYDGAAPRIAKGVIIIGN--GGSEFGARGFVSAPDAETGVDVRRFFLVPKPKEDPAASDS	25
QY	183	-----ETMGNDYEARMYGA-----WGQITTPYTNLVHYGSTAVGPASETOR	225
Db	255	VLMNKAYOTWS-----PGCAWTRGOGGGTYWSDIYVDPAVDLYLVLGVNGSPMNRYR	307
QY	226	GTPGCTLYGINTREAVRPTGETIWRHQTLPBNDMOECTFEWAVTYNDVQPSFEMGLQ	285
Db	308	SEKGDNLFLGISYALKPETGEYVWHFQETPRMDQDFTSDQIIMTLDPT-----	357
QY	286	SINPAATGGERVLVGPCKTGTMQFDEATEEFLMADNTYONKIBSID-ENGITYVNE	344
Db	358	-----NGETRHYIYHARKNGFFYIIDKTEEFISGKNYVYVWMAAGSLDPKRGRIYNE	410
QY	345	DALIKELDVEYDCPTFELGGRDMPASALNPDGSIYFIFLNVWCIDMAVDOEET-SMDY	403
Db	411	DALTTLGKRWYICIPDGLDGHNFAMAAISPKGLVYIYPAQOVPELYTNQVGFTPPBDSW	470
QY	404	NTS-----NVTKLPPG-----KDMIGRIDAIIDISTGRTLMSYVRAANYSPLISTGGGV	452
Db	471	NLGIDMKNKVGIPDSEPAKQAFYKDKMGITVAMDQOKAEAMRVDHKGPMWNGGLATGDDL	530
QY	453	LFNGGIDTRFRLASQETGETTLQOTFLATVASGOAISYEVDMQOYVAIAG	502
Db	531	LFQGLANGEFHAYDATNGSDLEHFAADGIIYAPRVTYLANGKROYAVENG	580

Db 681 EHTGKVVMSAVLPGGGQANPMTEANCHOYVALAGG 716

RESULT 5

US-07-642-734C-4

Sequence 4, Application US/07642734C

Patent No. 5824513

GENERAL INFORMATION:

APPLICANT: Donadio, S

APPLICANT: Katz, L

APPLICANT: Mcalpine, J B

TITLE OF INVENTION: Recombinant DNA Method for Producing

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Edward H. Gorman

STREET: Abbott Laboratories D377/Abpd-2 One Abbott

CITY: Park Rd

STATE: IL

COUNTRY: US

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/642,734C

FILING DATE: 17-JAN-91

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Danckers, Andreas M

REGISTRATION NUMBER: 32652

REFERENCE/DOCKET NUMBER: 4952.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 708-938-9396

TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3567 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-642-734C-4

Query Match 4.2%; Score 116; DB 2; Length 3567;
Best Local Similarity 23.4%; Pred. No. 0.55;
Matches 127; Conservative 47; Mismatches 232; Indels 136; Gaps 25;

21 ARAQVPTVDELINPAGEW-----ISYQONENRHSPLTQITENVGLOLW 71
Db 2911 ARAPIAVTDHRLARVDPG-MSDADAAPVIAV-----TTAHYALHDLAAGSGLIH 2964

72 ARGMPGKVVQVTPILHGVWYLANPGDVIAIDAKTGDLW-----EHRQLPNITATNS 126
Db 2965 AAAGVGMAAVALARRAGAEVLAAG-----PAKHGTLRALGLDDEHIASSRETFARK 3018

127 FGEPT--RGAIVAN--GVIVASTQYSPFGCFVS-GHDSATGELKRNFIIP-RAGE 179
Db 3019 FRERTGGRGVDVVLNSLIGELDDSSADLAEDEVEMGKTDLRADGDFRYAFPDIGE 3078

180 EGDETWMDY-EARWMTGANGQITYPVTNLV-----HYGSTAV--GPASE 222
Db 3079 AGDRDLGILREVYGLGA-GEIDRLPVSAMWELGSAPALQMSRGRHVKLVLPAPV 3137

223 TQRTG---PGTLVGTTRFAVRPDGELIWRHQL-PRDNDQECTEEMAVTNVQVPS 278
Db 3138 DPDCVTLITGT--GTIGRLAHVLTBEHGVRLILVSRGADAPGDELAEIEDIGAS 3195
QY 279 TEMEGLOSINPAAT-----GERRVLNGVCKTGTWMOFDAETGELIWRADTINYQMIESI 334

Db 3196 AETACDTRADNALSALDGLRPLTGVVHAAGVL----- 3230

QY 335 DENGIVT-VNEDAILKELDEVYDCPTFLGROMPSAALNPDSCITFIDLNVCYDMNAV 393

Db 3231 -ADGLVTSIDEPAVEQVYRAKVDAA-----WNLHELTANTGJSFVLFSSASVLA- 3280

QY 394 DQFTSMVYITSNV-----TKLPBGK-----DMIGRIDAID 425

Db 3281 ---GGQGVYAAANESINLAALRRTRGLPAKALGWLMAQASMTSGLDRIAR----- 3332

QY 426 ISTGRTLSVRAAANSPLYSTGGVLPNGGDRFRALSOETGETL-WQRLATVASG 484

Db 3333 --TGVAALPFRALALFDSALRRGGEVPEPLINSALRAAEFVEVLRGVRAKLRBAAG 3390

QY 485 QA 486

Db 3391 QA 3392

RESULT 6

US-08-439-009A-4

Sequence 4, Application US/08439009A

Patent No. 6004787

GENERAL INFORMATION:

APPLICANT: Donadio, S

APPLICANT: Katz, L

APPLICANT: Mcalpine, J B

TITLE OF INVENTION: Method of Directing Biosynthesis of

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Steven F. Weinstein

STREET: Abbott Laboratories D377/Abpd-2 One Abbott

CITY: Park Rd

STATE: IL

COUNTRY: US

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,009A

FILING DATE: 11-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Casuto, Dianne

REGISTRATION NUMBER: 40,943

REFERENCE/DOCKET NUMBER: 4952.US.D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-938-3137

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3567 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-439-009A-4

Query Match 4.2%; Score 116; DB 3; Length 3567;
Best Local Similarity 23.4%; Pred. No. 0.55;
Matches 127; Conservative 47; Mismatches 232; Indels 136; Gaps 25;

21 ARAQVPTVDELINPAGEW-----ISYQONENRHSPLTQITENVGLOLW 71
Db 2911 ARAPIAVTDHRLARVDPG-MSDADAAPVIAV-----TTAHYALHDLAAGSGLIH 2964

72 ARGMPGKVVQVTPILHGVWYLANPGDVIAIDAKTGDLW-----EHRQLPNITATNS 126
Db 2965 AAAGVGMAAVALARRAGAEVLAAG-----PAKHGTLRALGLDDEHIASSRETFARK 3018


```

1 LENGTH: 1042 amino acids
2 TYPE: amino acid
3 STRANDEDNESS:
4 TOPOLOGY: linear
5 MOLECULE TYPE: protein
6 US-08-928-361B-11
7
8 Query Match 4.1%; Score 113; DB 3; Length 1042;
9 Best Local Similarity 21.1%; Pred. No. 0.14;
10 Matches 139; Conservative 75; Mismatches 224; Indels 222; Gaps 36;
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
```

```

COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verry, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1837 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-5

```

```

Query Match      4.1%; Score 113; DB 3; Length 1837;
Best Local Similarity 21.1%; Pred. No. 0.36; Mismatches 224; Indels 222; Gaps 36;
Matches 139; Conservative 75;

```

```

QY 25 VPPVDELLANPAGWISYGNQ-----NYRHSPLTOITTEVNGOQLVWARGKOP 77
DB 1096 IDPTTG-LPFPNPTGHLINPTNNNTMDSFAGAYKIVASNGIKTDVNGLPVDITGLPK 1154
QY 78 GKQVTPPLIHGVMYLANPGDVIQAIDAKTGDLMEH-----RROLPIATLN-SFGE 129
DB 1155 DVSIDP-----FNSTTTEL-----VDPSTGKPINNTAGIVSGKRGILPDEBNGNLFD 1204
QY 130 PPRGMAIVANGVI-----VAGST-----COYSPGCFVSGSDATGEELMRNFIPRAG 178
DB 1205 PSTKLPIDGNQVLNVPETNSVSGSTSGTKPKR-GIPVNGGAVPDEE-----AK 1254
QY 179 EEGDETMGNDYEARMTGAMQIITYDPVNLVHYGSTA--VGPASETORGTPG-----229
DB 1255 DQADK--GKD--GLIYPPTNSINKDPVTNTOYSNNTGNIINP--ETGKVIYPSLPSLSLN 1307
QY 230 -----GLYGTNTRFAVAPDTGEIYWRHQTLPKRN--WDQECTFEKMAVTN 272
DB 1308 YPSFNTPOQDTGELTGKPDVDTYGLPYDPSTGEIIDPATKLPDGSVAGDEILLEVNLIT 1367
QY 273 -----VDVO--PSTEMEGLOS-----INP-----NATG-----ERRV 298
DB 1368 DEVTGLPIDLETGPRPVSGLPQLPNGTLVDSNKKPIPGSHSGFINSTSGSHKRP 1427
QY 299 LTGVPCKTGMWOFD-----AETGEFLMADTNYQNNI--ESTIDENGIVTWNEDAIL 348
DB 1428 STGKPLDPNPGGLPFDDESGSLINPETGDKLQSGHSGTFMPVPGKPOGEGGIMTPEO-LT 1486
QY 349 KELDY-----EYDVC-----FEL--GGGDMPSAL-----372
DB 1487 EALNKLPTSNVNIISPPSSDAVDPRTNTWMMKISGQYQVDDGKKTIPGSAASYIHIAL 1546
QY 373 -----NPDSGIYFIPLANNVCYDMAAVDQ-----EFTSMDYNTSNVTKLPP 413
DB 1547 GTPTQIDPTTGLPSPDSTGLPFIIPGNVVLVDPQTGEQIGSVYSLVYKKNITTEAY 1606
QY 414 GKDMIGRIDAIDISTGRITMSVRAAANYSPVLSG-----GVLFPNGGIDRIYFRA---464

```

```

DB 1607 GL-----PVDPKTGPIDPI-----SYLPFAKNGLIDIPISGYFSGSINAGFISGKAG 1654
QY 465 -----LSQFTGETLMQT-----RLATVASGAQISYEDVMQVAYLAG-----GGV 504
DB 1655 SQKSSDSDESNIPDPSTNMFYDKRTGKLIDPESGIAIDNSVSGV-FATVPTGAAPRKGGV 1713

RESULT 9
US-08-700-651-5
; Sequence 5, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUY, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4(HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1721
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-08-700-651-5

```

```

Query Match      3.9%; Score 107.5; DB 3; Length 1721;
Best Local Similarity 20.6%; Pred. No. 1.1;
Matches 137; Conservative 75; Mismatches 222; Indels 231; Gaps 37;

```

```

QY 25 VPPVDELLANPAGWISYGNQ-----NYRHSPLTOITTEVNGOQLVWARGKOP 77
DB 979 IDPTTG-LPFPNPTGHLINPTNNNTMDSFAGAYKIVASNGIKTDV-----YGLPV 1029
QY 78 GKQVTPPLIHGVMYLANPGDVI-----QADAKTGDL-----IMEHRROLPIATN 123
DB 1030 GEITGLP-----KPGSDIPFNSTGELVDPSTGKPLNNSTAGIYSGKRGLEPID 1080
QY 124 LN-SFGEPTRGMAIVANGVIASSTCOYSPGCFVSGSDAT-----GEELMRNFIP-- 175
DB 1081 ENGNLFPDSTLPLIDGNQVLNVPET-----NSTVSGSTSGTKPKRGPVINGGAVPDE 1134
QY 176 RAGEGDETMGNDYEARMTGAMQIITYDPVNLVHYGSTA--VGPASETORGTPG-----229
DB 1135 EAKDQADK--GKD--GLIYPPTNSINKDPVTNTOYSNNTGNIINP--ETGKVIYPSLPS 1187
QY 230 -----GLYGTNTRFAVAPDTGEIYWRHQTLPKRN--WDQECTFEKMA 269
DB 1188 SLNTPSNTPOQDTGELTGKPDVDTYGLPYDPSTGEIIDPATKLPDGSVAGDEILLEVNL 1247
QY 270 VTN-----VDVO--PSTEMEGLOS-----INP-----NATG-----E 295
DB 1248 ITTDEVYGLPIDLETGPRPVSGLPQLPNGTLVDSNKKPIPGSHSGFINSTSGSHKRP 1307
QY 296 RVLITGVPCKTG-MWOFD-----AETGEFLMADTNY-----Q 328
DB 1308 KDPSTGKPLDPNPGGLPFDDESGSLINPETGDKLQSGHSGTFMPVPGKPOGEGGIMTPE 1367
QY 329 NMIESIDENGIVTWN-----DAI-----KELDVEYVCPTFLIG--364
DB 1368 QILEAL--KILPTSNVNIISPPSSDAVDPRTNTWMMKISGQYQVDDGKKTITIGSASV 1425
QY 365 -----RDWPSAL--NPDSGIYFIPLANNVCYDMAAVDQ-----EFTSMDYNTSNV 408
DB 1426 IHTALGTPTQIDPTTGLPSPDSTGLPFIIPGNVVLVDPQTGEQIGKSVYSLVYKKNIT 1485

```


Query Match 3.8%; Score 105; DB 3; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.77;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 206 VTNIYHGSTAV-----GPAS-----ETORGPGTLYGTNTRF 239
 Db 1 MTNLMHQIIVPIRSLMPTGPASIPDDTEKHTLRSETSTYMLTVDGSGLIIVF 60
 QY 240 AVRPDGEIWRHQTLPD---NMDOCTEEMAVT-----NVDQPSFE 280
 Db 61 PGFP--GSVGAHTTLOSNGSYQDQ---MLTQNLDPVSYNGRLVRSRLTVRSSTL 113
 QY 281 MEGQSINP--NAATGRRVLTVGPCKGTMMQDPAETGEFLMARDTNYNNIEST--- 334
 Db 114 PGVYALNGTINAVT-----FQGSLEL---IDSYNGLSMTATNIN 152
 QY 335 DENGIYVNDALIKELDEYDVCPTFLGRMPASALMPDSGIYFIPLNVC----- 387
 Db 153 DKIGNVLVGGVYVLSLPTSLDYRLD-PIPAAGLDP-----KIMATCDSDRPR 204
 QY 388 -YDMAVADQ-EF-----TSMQVNTSNVTKLPKQMI-GRIDAIDISTGRTLM- 433
 Db 205 VYTVTADEXQFSSQLIPSGVKTTLTANIDALISLVGGELEFSCVTHIESTVDVTLYF 264
 QY 434 -----SYERAANYSPVLSGGVLFN--GGTDRFRALSOFTGETLMOTRLATVAS 483
 Db 265 IGDGTEVTVKAVADEGLTGTNNMLYFPNLGPGTSEITPITSMKLEVTYKKGCT--A 322
 QY 484 GOAISYEVDMQYVAIAG 502
 Db 323 GDPISWTSGLTAVTVIG 341

RESULT 12

US-08-931-608A-5
 ; Sequence 5, Application US/08931608A
 ; Patent No. 6302685

GENERAL INFORMATION:

APPLICANT: Sletat, David E.
 APPLICANT: Lobel, Peter
 TITLE OF INVENTION: NOVEL HUMAN LYSOSOMAL PROTEIN AND METHODS OF ITS USE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/931,608A

FILING DATE:

CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 601-1-077
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ. ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 635 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO
 US-08-931-608A-5

Query Match 3.8%; Score 104.5; DB 4; Length 635;
 Best Local Similarity 21.1%; Pred. No. 0.4;
 Matches 121; Conservative 65; Mismatches 239; Indels 149; Gaps 28;

QY 1 MKPTSLMASAGALIALAPAPAQVTPVDELLANPAGEWISYQNDQNYHSLTQIT 60
 Db 3 IKTALFVALIALMSSLSAIEDAMVSTHQAMSPASTQV-----LAASSTATT 54
 QY 61 TENNGQLQLYWANGMOGKQVOT---PLIHGVMIYLANPGVYQAIKAKTGDLIWEHR 116
 Db 55 TGNAYTILNMGSPRIDGAATVLEADHPLHVEALKLRNP-DALQTF----- 100
 QY 117 QLPNIATLNS--FGE-----PFRGM--AIYANGVVAOSTCYSPFGCFVSGH 160
 Db 101 -LAGVTTPGSALFKFLTFSQTERFPTQSOVDVAVHLOAQCFPTNIEVAPNRLLSAD 159
 QY 161 DSATGEELMKNYFIPRAGEGDETWGNDYEFARMTGAMGO-----ITYDPTNLVHY- 212
 Db 160 GTGAATNGFRFSIKRFSANGREFANDAPA-LVPASLGDSVNAVILGLQNSVAKHTLHHV 218
 QY 213 ----GSTAVGPASEFQRCPT-----GGTLIGVTPRAVRPDGEIYWRHQT 255
 Db 219 YHPEDVTVGPVNGVQAANAHAHPDPAIYGGSSLPAAFTNPAV---GITWGSIT- 273
 QY 256 PRDNDQECTFEMMYTNDVQPTMEGLQSNPNAATGERVRLVGPCKGTGMQDPAE 315
 Db 274 -----QTVVDLNSFTSGAGLATVNSTITKVSQGTFRANDPSNGE-WSLDSQ 318
 QY 316 -----TGEFLMARDTNYNNIESTIDENGIVT---VNEDALIKELDEYDVCPTFLG 363
 Db 319 DIVIAGVKKOLIFTSANGSSSGITDAGITASYRNAVTDNIAKILNSL----- 370
 QY 364 GRDWPASALN---PDSGIYFIPLNVCYDMMAVDOEFT---SMQYIN-TSNVTKLPQ 414
 Db 371 GEDETAAGQSGTQAADDAIF-----QAVVAGQGFSTIASGDAGVYQSTPTSGSRP 422
 QY 415 --KDMIGRIDALIDISTGFTMSVERAANYSP-VLSTGGVLFNGGIDTRFRALSOETGE 471
 Db 423 YVANSAGTV-KIDL---THYSVEPAS--SPYIVQVGGTTLTSGT-----TWSE 467
 QY 472 TLMOTRLATVASGOAISYEVDMQYVAIAGGVS 505
 Db 468 TVNNEGSLAIPASOG-----DNNQRLMATGGVS 496

RESULT 13

US-08-928-361B-30
 ; Sequence 30, Application US/08928361B
 ; Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETERS, VERNY, JONES & BIKSA
 STREET: 385 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA

COUNTRY: USA
 ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

$\frac{1}{2}$
 $\frac{1}{3}$
 $\frac{1}{4}$
 $\frac{1}{5}$
 $\frac{1}{6}$
 $\frac{1}{7}$
 $\frac{1}{8}$
 $\frac{1}{9}$
 $\frac{1}{10}$
 $\frac{1}{11}$
 $\frac{1}{12}$
 $\frac{1}{13}$
 $\frac{1}{14}$
 $\frac{1}{15}$
 $\frac{1}{16}$
 $\frac{1}{17}$
 $\frac{1}{18}$
 $\frac{1}{19}$
 $\frac{1}{20}$
 $\frac{1}{21}$
 $\frac{1}{22}$
 $\frac{1}{23}$
 $\frac{1}{24}$
 $\frac{1}{25}$
 $\frac{1}{26}$
 $\frac{1}{27}$
 $\frac{1}{28}$
 $\frac{1}{29}$
 $\frac{1}{30}$
 $\frac{1}{31}$
 $\frac{1}{32}$
 $\frac{1}{33}$
 $\frac{1}{34}$
 $\frac{1}{35}$
 $\frac{1}{36}$
 $\frac{1}{37}$
 $\frac{1}{38}$
 $\frac{1}{39}$
 $\frac{1}{40}$
 $\frac{1}{41}$
 $\frac{1}{42}$
 $\frac{1}{43}$
 $\frac{1}{44}$
 $\frac{1}{45}$
 $\frac{1}{46}$
 $\frac{1}{47}$
 $\frac{1}{48}$
 $\frac{1}{49}$
 $\frac{1}{50}$
 $\frac{1}{51}$
 $\frac{1}{52}$
 $\frac{1}{53}$
 $\frac{1}{54}$
 $\frac{1}{55}$
 $\frac{1}{56}$
 $\frac{1}{57}$
 $\frac{1}{58}$
 $\frac{1}{59}$
 $\frac{1}{60}$
 $\frac{1}{61}$
 $\frac{1}{62}$
 $\frac{1}{63}$
 $\frac{1}{64}$
 $\frac{1}{65}$
 $\frac{1}{66}$
 $\frac{1}{67}$
 $\frac{1}{68}$
 $\frac{1}{69}$
 $\frac{1}{70}$
 $\frac{1}{71}$
 $\frac{1}{72}$
 $\frac{1}{73}$
 $\frac{1}{74}$
 $\frac{1}{75}$
 $\frac{1}{76}$
 $\frac{1}{77}$
 $\frac{1}{78}$
 $\frac{1}{79}$
 $\frac{1}{80}$
 $\frac{1}{81}$
 $\frac{1}{82}$
 $\frac{1}{83}$
 $\frac{1}{84}$
 $\frac{1}{85}$
 $\frac{1}{86}$
 $\frac{1}{87}$
 $\frac{1}{88}$
 $\frac{1}{89}$
 $\frac{1}{90}$
 $\frac{1}{91}$
 $\frac{1}{92}$
 $\frac{1}{93}$
 $\frac{1}{94}$
 $\frac{1}{95}$
 $\frac{1}{96}$
 $\frac{1}{97}$
 $\frac{1}{98}$
 $\frac{1}{99}$
 $\frac{1}{100}$

Db 1188 SLNYPSENTPOQTDEITGKPV

RESULT 14
 US-08-928-361B-6
 Sequence 6, Application US/08928361B
 Patent No. 6071518
 GENERAL INFORMATION:
 APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 TITLE OF INVENTION: SPECIES INFECTIONS
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PETERS, VERNY, JONES & BIKSA
 STREET: 385 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION NUMBER: US/08/928,361B
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Verry, Hana
 REGISTRATION NUMBER: 30,518
 REFERENCE/DOCKET NUMBER: 480.76-1(HV)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1721 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-6

```

QY 270 VTN-----VDVQ--PSTEMEGLAS-----INP-----NATG-----E 295
Db 1248 ITTDEVTGLPDLTETGLPDPVSGSLPQLPGLTGLVDPSPNKKPIPGSHSGTNGTSGQSHE 1307
QY 296 RRVLTGVPCKTGT-MQOFD-----AETGEFLMARDNY-----Q 328
Db 1308 KDPSTGKPLPBNGLHFDHEDSGSLINPETGDKLQSHSGTETMPVPGKPGNGGIMPE 1367
QY 329 NMIESIDENGIVTNE-----DAI-----LKELDEYD-----V 357
Db 1368 QILEAL--NKLPTSNENINISPRSSDAVDPDRPTNWNKISQOTYQVDDKTTIPGSAASY 1425
QY 358 CPTFLG--GDMPSAAL--NPDGIYFIPLNNVCYDMAVDO-----EFTSMDEVNTSNV 408
Db 1426 IHTALGPTQDPTTGLPSPSTGTLPIPGFNVLPDQPGEQIKSVPLSVLYVEKIV 1485
QY 409 TKLPGKDMIGRIDAIDISTGRIMVEGAANYSFVSTG-----GCVLFNGSTDRYF 462
Db 1486 TEAYGL-----PDPPTGEPIDPI-----SYLPFAKNGLIDIPISGKYFSGS IAGFI 1533
QY 463 RA-----LSQETGELLMOI-----RLATVVASGOAISYVDGMOYVALAG----- 501
Db 1534 SGKAGSQSKSDSGSNPDPSTNMPYDRKGKLIIDPSGLAIDNSVSGV-FATVPTGAAP 1592
QY 502 --GGV 504
Db 1593 KKGCV 1597

RESULT 15
US-07-731-157A-4
Sequence 4, Application US/07731157A
Patent No. 5457032
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misset, Onno
APPLICANT: Van der Laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D., BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

```

```

MOLECULE TYPE: protein
US-07-731-157A-4

Query Match
Best Local Similarity 3.68; Score 98.5; DB 1; Length 816;
Matches 102; Conservative 73; Mismatches 178; Indels 199; Gaps 30;

QY 40 EWISYGCNDENRHSPLTQITTEENGOLOLVNARGM-----QKQVQVPLHHDGMITLAN 95
Db 145 EYVDF-----DFQPEPLTDF-----DYMIWGSKANRFSOTNIEVALA----- 184
QY 96 PGYVAIDAKTG-----DLW-----EHRRLPNITLNSFGEP 131
Db 185 --WROSLKORHGPERRALFDELLIMDPTAPTTPAPAEHKKQAO-----AGT 232
QY 132 RGMATVANGVIYAGSCQYSPGCFVSGHDSA-TGEELMRNYFIPRAGEGDETFWGDYE 190
Db 233 QDLAHVSSPYLATTELRQDKHMG--GRGPDAPAKSNLMSTR--PERVQEGSTVLINGPQ 288
QY 191 ARWMTGAMQOITYDPVTLNVHGSTAVGPASBETORGTGPG--TLXGNTFRFAVRPTGE 247
Db 289 FGV-----YNDAYT--YGIGLHAGGFDVYGNTPFAIPVLFGTNS-----E 327
QY 248 IYWRHQTLPROMWDOECTFEMAVTNVDVOPSTEMEGLQSNPNA-----T 293
Db 328 IAMGATAGPOD-----VVDIYQ-----EKLNPSSADQYWFNNMARTMEQR 367
QY 294 GRRVLTGVPCKTGTMM-----QPDATGE-----FLMARDNYON 329
Db 368 KERIOVRCQADREMTIMTWTGHPVMOEDYDGAAYSKRSMWDGYEVOQLAMLVAKARN 427
QY 330 MIESIDENG--IVYVNEDEALIKELDEYDYVCPFL-----GGRDMP-- 369
Db 428 WTEFLDQASKMAISINWYADKHGNIGY-VSPATLPQRPADODIRVPKADGSMELGK 486
QY 370 -----AALNDSGIYFIPLNNVCYDMAVDOETSMDEVNTSNVTKLPKXDMIGRIDA 423
Db 487 SFDAIPKAYNPPOG-YLVNWN-----KPADK--TNTDTYTWYGGRM--NELVSQYQO 536
QY 424 IDISTGRTMSEVERAANYSFVSTGGCVLFNGGTDRYFALSOEFGTLMOTRLATVAS 483
Db 537 KDLFSVQEIWEFNO-KASYDY-----NMKTFRPHLEKLAQOLPADDSKAL 583
QY 484 GOAISYEVDGMO 495
Db 584 TMLLAW--DGME 593

```

Search completed: May 24, 2002, 10:20:18
Job time: 331 sec

•
• Fri May 24 11:27:32 2002

walick-934-135.pep.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:21:08 : Search time 54.76 Seconds
(Without alignments)
898.425 Million cell updates/sec

Title: WALICK-934-135.PEP

Sequence: 1 MKPTSLMMSAGALALIAAP.....GMQYVAIAGGVSYGSLINS 512

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:1:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	18.1	742	A49340	alcohol dehydrogen
2	487	17.8	601	E95863	alcohol dehydrogen
3	485.5	17.7	708	E52317	quinohemoprotein e
4	483.5	17.6	738	S14270	alcohol dehydrogen
5	483	17.6	742	JS0326	alcohol dehydrogen
6	434.5	15.8	626	JQ0706	alcohol dehydrogen
7	417	15.2	623	B83399	quinoprotein alcon
8	392.5	14.3	573	S68591	methanol dehydroge
9	349.5	12.7	796	JY0107	glucose dehydrogen
10	349.5	12.7	796	H90644	glucose dehydrogen
11	349.5	12.7	796	H85495	glucose dehydrogen
12	340.5	12.4	796	AG0523	glucose dehydrogen
13	329	12.0	801	S00943	glucose dehydrogen
14	294.5	10.7	808	1 OPEKX	glucose dehydrogen
15	279	10.2	803	F83360	glucose dehydrogen
16	276	10.1	809	B98314	hypothetical prote
17	276	10.1	809	A12968	hypothetical prote
18	268	9.8	778	G98221	glucose dehydrogen
19	268	9.8	778	A13064	glucose dehydrogen
20	254.5	9.3	809	A55547	quinone-shikimate
21	220	8.0	639	2 JC4881	polyvinyl-alcohol
22	209.5	7.6	521	A41378	hypothetical prote
23	157.5	5.7	524	A82580	polyvinylalcohol d
24	144.5	5.3	525	A99953	hypothetical prote
25	142.5	5.2	392	AG0821	probable lipoprote
26	142	5.2	839	D97013	probable secreted
27	140.5	5.1	668	C75264	probable serine/th
28	137.5	5.0	392	F91050	probable dehydroge
29	137.5	5.0	392	B85895	probable dehydroge

30	136.5	5.0	392	2	G65027	hypothetical prote
31	134.5	4.9	407	2	H69064	serine/threonine p
32	132.5	4.8	386	2	A82284	conserved hypotet
33	131	4.8	2535	2	AC0304	probable hemolysin
34	126	4.6	407	2	C82804	conserved hypotet
35	125.5	4.6	827	2	F64512	hypothetical prote
36	120.5	4.4	3573	2	S23070	erythronolide synt
37	119.5	4.4	1588	2	A86036	probable adhesin B
38	119.5	4.4	1588	2	H91188	WD-40 repeat prote
39	119	4.3	1526	2	AC2239	serine/threonine p
40	118.5	4.3	799	2	T48889	hypothetical prote
41	117.5	4.3	1197	2	D82696	Rhs-family protein
42	117	4.3	1354	2	AG0538	regulatory protein
43	115.5	4.2	954	2	A87431	outer membrane por
44	115	4.2	443	2	S23771	coatomer complex D
45	114	4.2	796	2	T39962	

ALIGNMENTS

RESULT 1
A49340 alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113)
C:Species: Acetobacter pasteurianus
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Cross-reference: A49340, MID:94042848
A:Reference number: A49340
A:Accession: A49340
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-742 <TAK>
A:Superfamily: glucose dehydrogenase (pyridoxin-dependent quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match	18.1%	Score 495	DB 2	Length 742
Best Local Similarity	27.0%	Pred. No. 8.9e-28		
Matches 164	Conservative 82	Mismatches 221	Indels 140	Gaps 22
QY	11	AGATALLAAPFAOVTPYTDLL--ANPPAGEWISYQONENYRHSPLQITTEYNGQD 68		
DB	23	AAALPYAAVPPARADGQGTGAIHADDHPENMLSYGRYSQRYSPLDQINRSNVDLK 82		
QY	69	IWMARGMPGRVQ-VTPLIHDGVMYLANPGDVIOAIDAKGDLIWEHRQLP-NIA---- 122		
DB	83	LAWYTLDTNRQENTPLVVDGIMYATTNMSKMDALDAATGKLLMDYDPKPGNIADKGC 142		
QY	123	--TLN-----SFG-----PTRGMAT 136		
DB	143	CDTVNRGAGYNGKVFQWGTFGDGRVLAADAKTGKWEVNTIPADASIGKRSYVVDGAVR 202		
QY	137	VANGVIVAGSTOYSPFC--FVSGHDSATGEBELRNFTIRAEED----- 182		
DB	203	VAKGVLIGN--GSESEFARGFVSADFETGKLRKFTVNNKNEPDHVAADVNLMSKA 260		
QY	183	-ETWGDYDARKMT-----GAMGQITVDPVTVNLVYGSTAVGPASETORGPGLTY-- 233		
DB	261	YKTMGP--KGAVRQGGGTWDSLVYDPSDLIY--LAAG-----NGSPMYKYSSE 309		
QY	234	--GTN---TRFAVRPDTEIIVRHQTLPRNDWDOCEFFEMVTVNDVQPTENMGQSI 287		
DB	310	GIGSNFLGISYVALKRETEGYVWHFOATPMDQMDTSVOQITFTLDMPV----- 357		
QY	288	NPAALAGE--RRLVTPCKTGTMQDAGTGEELNARDNYNNMIESIDE--NGIYTVNED 345		
DB	358	-----NGEMRHVIMAP--KNGFTYVDAKTGELAGKNVYNNMANGDPLGRIPYPD 411		
QY	346	AIKELDVEYDVCPFTLGGHDPMSAALNPDGIIYPLINNVCY-----DMMA 392		

Db 412 GLYTLGKFWYIGIPGLAHNFMGMAYSPTKHLVYLPDAHPIPCYKNOVGCFPHDPANN 471
 QY 393 VDOEFTSDVNTSNVTKLPKGMIGRIDALIDISTGRITLMSVERAANYSPVLTGGGV 452
 Db 472 VGLDMTKGLDTPDE-ARTAYIKDLHEMLAMPVKKEFTWKIDHGMWNGVLTATGSDL 530
 QY 453 LFNCGTDYFRLASQETGETLMOTRLATVASGOAISYVDGMQYVA-----IAGG 502
 Db 531 LFOGLANGEEFAHYDATNGSDLYKFKDAQSGIIASPMYTSVNGKOYVAEVEGMIPIPSMG 590
 QY 503 GVSTGSG 509
 Db 591 GVGRTSG 597

RESULT 2

alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - *Sinorhizobium*
 C:Species: *Sinorhizobium meliloti*
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
 C:Accession: E95863
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A>Title: The complete sequence of the 1,683-kb *psymb* megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95863
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-601 <K0>
 A:Cross-references: GB:AL591985; PIDN:CA048573.1; PID:915140045; GSPDB:GNO0167
 A:Experimental source: strain 1021, megaplasmid *psymb*
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abolia, P.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 Hubbard, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C:Genetics:
 A:Gene: *Smb20173*
 A:Genome: plasmid
 C:Keywords: oxidoreductase

Query Match 17.8%; Score 487; DB 2; Length 601;
 Best Local Similarity 26.2%; Pred. No. 2, 5e-27;
 Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

QY 14 LALLAAPAFQAVTPYDEL--LANPAGWISYGOENYRHSPLTOITTEVNGOLQVW 71
 Db 8 LAISIGGAGVAFANDELQKIDDP-NQWAIOTGYANIKRSKLDQINKDNVQKIQVAM 66
 QY 72 --ARGMQPGVQVTPPLIHGCVMTLANP-GDVIOAID-ARTGDLIMHR-RQLPNATLNS 126
 Db 67 TFSGTGVRGH-EGSPPLVIGDLMVHTPFVNTYALDLSKQGVYMKYKQDPNVIPIWC 125
 QY 127 FGEPTRGMAI-----VANGYIVAGST----- 147
 Db 126 CDIVNGVAADAKIFLHQADITVVALDAKTKVYIWSVKNNGATKGTNTATVMPVKIKI 185
 QY 148 ---CQSPFGC--FVSGHDSANGELMRYFI-----PAGEESDET 184
 Db 186 LVGISGEGVGRVHTAYASMDQKVLKRGYSKGPDSDTLIDEKTHLGKPYGKDSGLTT 245
 QY 185 WGNDFYARMWTG--AWGIIYDPVTLNLYHGSTAVGPASETFQRTPGGTLVGTNTRAV 241
 Db 246 WEED---QMKIGGCTTGWYSTDPEENLYYGTGNPSTWNPOR--PDGNBW-SWTIPAR 299
 QY 242 RPDGTGIWRHQTLPDNDQCTEEMAVTNVDQPTSTEMGLSINMNAITGERRVLTG 301

Db 300 DVDTGMAKWLQYMPHDEWIDYGVNEMILTEQOJDGK-----DRKLTH 343
 QY 302 VPCKTGIMQPDATGTEFL-----WARTNQNMI-----SIDENGIV 340
 Db 344 FD-RNGGYTDRTYTGELVAEKYDPTVWNAETEVDMPKSDKGRPOVQAQSTFONG-- 400
 QY 341 TVNEDAILKEIDVYDCPTFLGRDMPSAALNPDSEIFIFILNNVCYMAVNDDEFISM 400
 Db 401 ---EDTWT-----GVCFAALGTDKQOPRAYSFKTELTFVFNHYCNDYEPFRVSYTAG 451
 QY 401 DVYNTSNVTKLPKGMK---IGRIDALIDISTGRITLMSVERAANYSPVLTGGVFNMG 457
 Db 452 QPIYGAATLSWYP-KDSHGGMGNFIMDNKEKIKWSLPEPFSVSGALATAGDVVYCT 510
 QY 458 TDRYFRLASQETGETLMOTRLATVASGOAISYVDGMQYVAAG-----GVSTGSGINS 512
 Db 511 LEGYLAHVDAATGKELKRYKRTSPGVGNVYIAREGKOYVAVLSGVGMAGITGLAAGLTN 570

RESULT 3

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - *Comamonas tes*
 C:Species: *Comamonas testosteroni*
 C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S62366; S62373; S65908; S52317
 R:Stoorvogel, J.; Krayveld, D.E.; van Sijts, C.A.; Jongejan, J.A.; de Vries, S.; Dui
 Eur. J. Biochem. 235, 690-698, 1996
 A>Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenas
 A:Reference number: S62366; MUID:96184549
 A:Accession: S62366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STO1>
 A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
 A:Accession: S62373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STO2>
 R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Dui
 Eur. J. Biochem. 230, 899-905, 1995
 A>Title: Quinohemoprotein ethanol dehydrogenase from *Comamonas testosteroni*. Purific
 A:Reference number: S65908; MUID:95324580
 A:Accession: S65908
 A:Molecule type: protein
 A:Residues: 32-50; 'X', 52-54; 477-483, 'X', 485-490 <DEU>
 A:Experimental source: ATCC 15667
 C:Genetics:
 A:Gene: *qhcdh*
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; quinoprotein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 17.7%; Score 485.5; DB 2; Length 708;
 Best Local Similarity 25.7%; Pred. No. 4e-27;
 Matches 152; Conservative 86; Mismatches 241; Indels 113; Gaps 20;

QY 3 PSLIMASAGALALLAAPAVTPY-----DELANP-AGENISYGOEN 50
 Db 12 PGRMWTLLAACLG--SAAPFQGTGPAQAALAAVORVDDDFRANARRPDPITGVYAE 69
 QY 51 YRHSPLTQITTEVNGOLQVMAKQPGK-VQVTPPLIHGCVMTLANP-GDVIOAIDARTGD 109
 Db 70 TRYSRLDQIANAVKDLGLAWSYNLESTRGVETATPVVDGIMYSASVYHAIDITRGN 129
 QY 110 LIMEHRRQIP-----NIAT-----LNSFG 128
 Db 130 RIWITDPOIDRSTGKGCDDVVRGVALMKGVYVAGWDRLIALDATGKEVHWQTF- 188
 QY 129 EPTRGMAIYVANG-----VIVAGSTCOYSPFGCFVSGHDSANGELMRYFI----- 175
 Db 189 EGOKGSLITITGAPRVFKKGVIIKRGAEYGRG-YITAYDAETGERKRWFSVPGDSPK 247

QY 288 NENATGE-RRLVLTGVPCKTGTMMQFPAETGEFLARDNTYQNMIESID-----EN 337
 Db 359 -----GEMRHVYHAP-KNGFFVLDLAKTGEPLSGNINYYQWANGCLDPLGRPMNDP 411
 QY 338 GIIVYNEALIKELDEYDVCPTFLGRDWPSSALNDGSIYPIPLNNVCY-----388
 Db 412 GLVTLNG-----KFWYGI-PGFLGAHNFMAAAYSPKTHLYIIPAHQIPFGYKNGVGGE 463
 QY 389 -----DMAVDOFTSMQVNTSNTKLPKGDMDIGRIDALDISTGRTLMSVERAANYSF 444
 Db 464 KPHADSNVNGVDMTKKNLPTPE-ARKAYIKDLHGWLADPYKMETVWKIDHKGPMNGC 522
 QY 445 VLSGGCVLLENGTDYFRALSOETGETIMQTRLATVASSQAISYEDMOQYVA-----498
 Db 523 ILATGDLLEFQGLANGEFHAYDATNGSDLYKFDAGSIIAPPMYTSVNGKQYVAIVEVNG 582
 QY 499 -----IAGGVSYSG 509
 Db 583 GIYPSMGVGRISG 597

RESULT 6

JQ0706
 alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
 N:Alternate names: methanol dehydrogenase 62K large chain
 C:Species: Methylobacterium extorquens
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
 R:Accession: JQ0706; S07908
 R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
 Gene 90, 173-176, 1990
 A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 moxJ and moxJ genes
 A:Reference number: JQ0706; MUID:90337342
 A:Accession: JQ0706
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-626 <AND>
 A:Cross-references: GB:M31108; NID:g150017; PIDN:AAA2380.1; PID:g150018
 A:Experimental source: strain AM1
 R:Nunn, D.J.; Day, D.; Anthony, C.
 Biochem. J. 260, 857-862, 1989
 A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
 A:Reference number: S04644; MUID:89350892
 A:Accession: S07908
 A:Molecule type: protein
 A:Residues: 28-50, 'XX', 53 <NUN>
 A>Note: The source is designated as Methylobacterium extorquens AM1
 C:Comment: This enzyme oxidizes methanol to formaldehyde.
 C:Genetics:
 A:Gene: moxJ
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match 15.88; Score 434.5; DB 2; Length 626;
 Best Local Similarity 26.33; Pred. No. 1.6e-23;
 Matches 156; Conservative 91; Mismatches 219; Indels 127; Gaps 26;
 QY 10 SAGALALLA-AP-AFAQVTPVTDPELLANPAGG-WISYQONQNYRHSPILQITTEVQ 66
 Db 7 SVSLAMALALPALSSGAYANXKLVLSKSDNWMVPGKNDYNNFSLDKOIKNGNVK 66
 QY 67 LQLVN-ARGMQPGVQVPLIHGVVY-----LANPGDVY-----100
 Db 67 LRPAMTSTGLNGH-EGAPLVVDGKMYIHSPNNTPALGLDDETTIMQKPKQNPAA 125
 QY 101 QAI-----DAKTGLIWEHRRLP-NIATLNSGFEPTRGMAIVANGVYA 144
 Db 126 RAAVACCDLVNRGLAYWPGDKTPLLIL-KTQLDGNVALNA-ETGEIVMKVEHSDIKV 181
 QY 145 GSFQVSPF-----GC-----FVSGDSTGELMKNRFT-----174
 Db 182 GSTLTITAPYVVKKVIIGSSGAEVGLVLTAVDVTKGBOVRAVATGPDKDLIASDFN 241

QY 175 ---PRAGEE--GDETWGNDYEARMITGA---WQOITYDPVTLNVHGSTAVGAPASETORG 226
 Db 242 IKNHYQKGKLGITWEGD---AMKIGGTNMGWYAYDETNLITHTGNPAPMETMK- 297
 QY 227 TPGGTLYGTNTRFAPVDPGEIYWRHQTLEPRDMWDECTEFEMVTVNVDQSTEMEGLQS 286
 Db 298 -PGDNKM-TWITIGRDAITGEAKFGYQKTPHEMD-----YAGVNNMLSEKD--- 344
 QY 287 INPAAATGERRVLTGVPCKTGTMMQFPAETGEFLAR---DTNYQNMIESIDENGIYVN 343
 Db 345 -----KDGARKLLIHPDRNGIYVTLDRDGLVSNKRLDPT--VNVKSDVLKTGPVR 397
 QY 344 EDALIKELD-VEYDVCPTFLGRDWPSSALNDGSIYPIPLNNVCYDMAVDOFTSMV 402
 Db 398 DPEYGTMDHLADICPSAAGYHNGHSDYDPKRELFFNGIHICMDHEPMLPRBAQF 457
 QY 403 YNTSNTKLPKPK-----DMIGRIDALDISTGRTLMSVERAANYSFVLSGGCVLENG 456
 Db 458 FVGATLMMTPGPKGRONTEGLQIKAYNAITGDYKWKEMERFAVWGTMATAGDLVFG 517
 QY 457 GTDYEFRALSOETGETIMQTRLATVASSQAISYEDMOQYVAIAGGVSYSG 509
 Db 518 TLGGYLKARDSDTGLDMKFKIPSGALIGYPMYTHKGTQYVAI-----YIGVG 565

RESULT 7

B83399
 quonoprotein alcohol dehydrogenase PA1982 [Imported] - Pseudomonas aeruginosa (strain
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: B83399
 R:Stoyer, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: B83399
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-623 <STO>
 A:Cross-references: GB:AE004624; GB:AE004091; NID:g9947973; PIDN:AA05370.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: exaA; PA1982

Query Match 15.24; Score 417; DB 2; Length 623;
 Best Local Similarity 23.18; Pred. No. 3e-22;
 Matches 140; Conservative 95; Mismatches 247; Indels 124; Gaps 20;
 QY 3 PTLMSASGALAL---LAPAPQVTPVTDPELLAN--PPAGWISYQONQNYRHSPIL 57
 Db 9 PAGLIRPSHLCLAPVALVLSGAGALAKADVWEDIANDDTKTGVDLYQGTGAQMSPLK 68
 QY 58 QITTEVNGOLQVWARGM---OPGKQVTPVPLIHGVVYLANGVYQIADKTTGDLTME 113
 Db 69 QVADNVETKLLPMSYSFGEDEKRGQ-ESQATYSDEVITYVTASYSKLPALDAKTKRLMT 127
 QY 114 HRRLP-----NIATLNS-----126
 Db 128 YNHLPPDIRPCDDVVRGAAIYDKVFEGLDASVALNKMTGKVVYMKKRPADHAGYT 187
 QY 127 -FGEPTGMAIVANG-----VIYAGSTQCYSPGCC--VSGDSTGELMKNRFTIPR- 176
 Db 188 MTGAPT-----IYADGKTGKVLTHGSS--GDEFVYGRFLFARDPDTGEBIWRPVEGSHM 241
 QY 177 -----AGEGDETWGNDYE-----RMNTGA---WQOITYDPVTLNVHGSTAV 217
 Db 242 GRNGKDSYTVGVKAPSNPDDRNSTPKVESMSHGAGPQWOSAFDEETITIIYVAGNP 301
 QY 218 GPASEYRGKTPGG-----TLIGTNTFAVRDPDTEIYWRHQTLEPRDMWDECTEFEMVTV 271

```

Db 302 GPMWTAARAKAGNDHDSLT-TSGQGVDPSPSEVWFQHPNPDAWPFSGNNELVLF 360
OY 272 MNDVDPSTMEGLQSTINP-----AATGRRYLGVPCKTGTMT--QEDATGTFEWARD 324
Db 361 DYAKAGKLVKATATAHADRNGFFVYVDRSNGKLNAPFVDNITWASHDLTG-----RP 415
OY 325 TYNQNMIESIDENGLVTVMEDALIKELDEYDVCTPLFGSDWPSAALNPDGTYEPLN 384
Db 416 VEREGORPPLPEPG-----OKHKAAYVSPPLFGKNNWNNMAYSDQTLFYPVAN 465
OY 385 NVCYDMAVAQDEFTSMQVNTSNVTKLPKCKMIGRIDALIDISTRTTWSERAANSP 444
Db 466 HMKEDVTEEVSYTKGSAVLGMCGRFRKMYDDHVGSLRAMPVSGKVVWEKHEHPLWAG 525
OY 445 VLTGGGVLENGTIDRYFRALDSQETGELTMQTRLATVASGAISYEVMQVYALAGGV 504
Db 526 VLTAGNLVFTGTGDYFRAFAKSKGELMKFOTGSGIVSPITWEOGEOYLGVT---V 582
OY 505 SYSGSL 510
Db 583 GYGAV 588

```

```

RESULT 8
S68591
methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylotrophus methylotrophus (strain
C:Species: Methylotrophus methylotrophus
A:Variety: strain W311
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S68591
R: Xia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.
J. Mol. Biol. 259, 480-501, 1996
A:Title: Determination of the gene sequence and the three-dimensional structure at 2.4 Å
A:Reference number: S68591; MUID:96256524
A:Accession: S68591
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-573 <X1A>
A:Cross-references: EMBL:U01040; NID:g1127819; PIDN:AAA83765.1; PID:g1127820
A:Note: the authors did not translate the codons for residues 1-2
C:Keywords: oxidoreductase

```

```

Query Match 14.3% Score 392.5; DB 2; Length 573;
Best Local Similarity 24.6% Pred. No. 1.6e-20;
Matches 136; Conservative 84; Mismatches 218; Indels 115; Gaps 22;

OY 38 AGEMISYGNQENYRHSPLTQTTENWGLQVW--ARGMQPKVQVYPLIHGCVMTL-- 93
Db 13 AGAWPLATGGYSSQNSPLAOLKNSVKNVKAWSFTGVINGH-EGAPLIVIGDMYVHS 71
OY 94 ANGVYQAIQAKTDLIWEHR-RQLPNIALNLSGFEETRGMA-----IV---ANGVYA 144
Db 72 AFNNNTYALNLDPKRIWQHRPKDASTKAVMCCDVYDRGLAAGAGIYVKKANGHLLA 131
OY 145 -----GSTQYSPF-----GC-----FVSGHDSATGEBEL 168
Db 132 LDAKTKIMWEYVCDPKKGSITLQAFYAKQTVLMGSGAELGVAVNAFDTKGTBLK 191
OY 169 WR-----NYFLPRAGE--EGDETGNDEYEARWNGA---WGQITVDPVT 207
Db 192 WRAFATGSDSVRLAKDFNSANPHYGQFLGRTWEGD---AMKIGGCTNGWYAYDPKL 248
OY 208 NLYHVSSTAVGPASETQRTGGPG-----TLVGTNTRFARVPDTGEIWRHQTLPDNDQ 262
Db 249 NLYYSGGNPAPWNETMR--PGDNKRYMTIWRDL-----DTGAKKGYQKTPHDEWD 300
OY 263 ECTFEEMVTVNDVQSTMEGLQSTINPNAATGERRYLTVPCKTGTMMQFADTEGEFLMA 322
Db 301 AGVQWAVLTD-----QPVN-----GKMPPLSHIDRNGILTYTLNRENGNLIVA 343
OY 323 RDTN-YQMIESIDENGLVTVMEDALIKELDVE-YDVCTPLFGSDWPSAALNPDGTYE 380

```

```

Db 344 EKVDAVAVVEKVKDLTKGTGTPVDRPEFATRMHKGNTNLCPSAMGFHNOGVDSYDSESTTK 403
OY 381 IFLNNVCYDMAVAQDEFTSMQVNTSNVTKLP-----PECKMIGRIDALIDISTGRTLMSVE 436
Db 404 AGLNICHMDEPFMLPYRAGOFVGAATLAMYPPGNPLPKKMGQIRAFDLITGAKAKTKW 463
OY 437 RAANSPVLTSGGVLENGTIDRYFRALDSQETGELTMQTRLATVASGAQISYVDGMQY 496
Db 464 EKFAWAGTLTKGGLVYATIDGLKALDNKDKGELMNFMPKPSGGIGSPMTYSFKKQY 523
OY 497 VALAGGVSYSGS 509
Db 524 I-----GSMYGVG 531

```

```

RESULT 9
JV0107
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli
C:Species: Escherichia coli
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64735; JV0107; A45997; S45201; I41228
R: Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <LNU>
A:Cross-references: GB:X51323; NID:q1786315; PIDN:AMC73235.1; PID:g17863
A:Experimental source: strain K-12, substrain MG1655
R: Cleton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino
A:Reference number: JV0107; MUID:91035240
A:Accession: JV0107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51323; NID:q41553; PIDN:CAA35706.1; PID:g41554
A:Experimental source: strain K12
R: Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Topological analysis of quino protein glucose dehydrogenase in Escherichia co
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R: Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <FNU>
A:Cross-references: EMBL:D26562; NID:q473770; PIDN:BA00580.1; PID:g473791
A:Experimental source: strain K-12 substrain W3110
R: Tamada, M.; Asaka, S.; Sailer, M.H.; Yamada, Y.
J. Bacteriol. 175, 568-571, 1993
A:Title: Characterization of the gsd gene from Escherichia coli K-12 W3110 and regula
A:Reference number: I41228; MUID:93123180
A:Accession: I41228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
A:Cross-references: GB:D12651; NID:g21655; PIDN:BA002174.1; PID:g216556
C:Genetics:
A:Gene: gsd
A:Map position: 3 min
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain

```

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:11-37/Domain: transmembrane #status predicted <TM1>
 F:41-59/Domain: transmembrane #status predicted <TM2>
 F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:466/Active site: Asp #status predicted

Query Match 12.7%; Score 349.5; DB 1; Length 796;
 Best Local Similarity 22.8%; Pred. No. 3.3e-17;

Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;

```

QY 6 LMAS-----AGALALLAFAQVPTVDLLANPAGWISYGCNENYRHSPLTQ 58
DB 135 LTMAGFNDPQEIINGTISADATPAEA-ISPVAQ-----DMPAYGRNCGORFSPKQ 185
QY 59 ITTENVGOLQVWA-----RGMPGRV--QVTPLIHDGVMYLANPGDVIQAIQAKTGD 109
DB 186 INADVHNHKEAWVFRTGDVQKQPNDEGITTVEPIKVGDTLYICTAHORLFDALDASGK 245
QY 110 LIMEHRQLPNTATLNSFG-----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
DB 246 EKMHYD--PELKTNEFQHTVCRGVSYHEAKAETASPEVADCPRIILPVDGRLIAI 302
QY 129 -----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
DB 303 NAENGLKETPANKGVNLQSNMPTDKPGLYEPTSPRIITDKTIVMAGSYTDFNSTRETS 362
QY 155 CFSVGHSDATGEBELRNFEIPRA-----GEEDDETGNDEYARMWTGAMQITDYPVT 207
DB 363 GVIRFEDVNTGELLMA--FDGAKDPNAPISDEHTFTENS-----PNSMAPAAIADAL 413
QY 208 NLVHGSTAVGPASSETQGTGTLGTNRFAVRPDGTGEIVRROTLPDNDMDQECTFE 267
DB 414 DLVYLPKGVTPPDIMGKRTPEQERIASI-LALNATTGKLAWSYQTVHNDLMDMD-----468
QY 268 MMVTNVDOVPSSTENEGLOSINPNAATGERVLTGVPCKTGMQFDETEGFL-----320
DB 469 -----LPAQPT-----LADITVN--GQKVPVIYAPAKTGINIFVDRRNGELVVAPEK 515
QY 321 ---WARDTNGNMIESIDENGI-----VTVNDAIILKEL--DVEYD-----356
DB 516 VPQGAAGDVTPTQPPSELSFRPTKDLGADMMGATMFDQVCRVMEFHOMRYEGITPP 575
QY 357 -----VCPFTLGRDMPSPAALNP-----374
DB 576 SEQGTLVFPNGLMFEWGGISVDPNREVAIANPMALPFSKILIPRGPNMEQPKDAKT 635
QY 375 ---DSGI-----YFIPLNVCYDMAVADQETSMQVNTSNTKLPKPKDMIGRIDAID 425
DB 636 GTESGIQPOYGVPTVTLN-----PFLS-----PFGLPCKOPAMGVIYSLAD 676
QY 426 ISTGRITMSVERAANYS-----PV-----LSTGGVLENGGT--DRYPRAL 465
DB 677 LKTNEVVMKKRIQTPDQSMFPMPVPVFNMGMPMLGPISTAGNVLEFIATADNYLRAV 736
QY 466 SOETGETLMOQLATVYASGA--ISYEDGMQYAI-AGGVSYGS 508
DB 737 NMSGKELWGRLP--AGQATPMTEYVNGKQIVVISAAGHSFST 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A:Reference number: A99629; MID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA93551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: Rcs0128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.7%; Score 349.5; DB 2; Length 796;
 Best Local Similarity 22.8%; Pred. No. 3.3e-17;

Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;

```

QY 6 LMAS-----AGALALLAFAQVPTVDLLANPAGWISYGCNENYRHSPLTQ 58
DB 135 LTMAGFNDPQEIINGTISADATPAEA-ISPVAQ-----DMPAYGRNCGORFSPKQ 185
QY 59 ITTENVGOLQVWA-----RGMPGRV--QVTPLIHDGVMYLANPGDVIQAIQAKTGD 109
DB 186 INADVHNHKEAWVFRTGDVQKQPNDEGITTVEPIKVGDTLYICTAHORLFDALDASGK 245
QY 110 LIMEHRQLPNTATLNSFG-----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
DB 246 EKMHYD--PELKTNEFQHTVCRGVSYHEAKAETASPEVADCPRIILPVDGRLIAI 302
QY 129 -----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
DB 303 NAENGLKETPANKGVNLQSNMPTDKPGLYEPTSPRIITDKTIVMAGSYTDFNSTRETS 362
QY 155 CFSVGHSDATGEBELRNFEIPRA-----GEEDDETGNDEYARMWTGAMQITDYPVT 207
DB 363 GVIRFEDVNTGELLMA--FDGAKDPNAPISDEHTFTENS-----PNSMAPAAIADAL 413
QY 208 NLVHGSTAVGPASSETQGTGTLGTNRFAVRPDGTGEIVRROTLPDNDMDQECTFE 267
DB 414 DLVYLPKGVTPPDIMGKRTPEQERIASI-LALNATTGKLAWSYQTVHNDLMDMD-----468
QY 268 MMVTNVDOVPSSTENEGLOSINPNAATGERVLTGVPCKTGMQFDETEGFL-----320
DB 469 -----LPAQPT-----LADITVN--GQKVPVIYAPAKTGINIFVDRRNGELVVAPEK 515
QY 321 ---WARDTNGNMIESIDENGI-----VTVNDAIILKEL--DVEYD-----356
DB 516 VPQGAAGDVTPTQPPSELSFRPTKDLGADMMGATMFDQVCRVMEFHOMRYEGITPP 575
QY 357 -----VCPFTLGRDMPSPAALNP-----374
DB 576 SEQGTLVFPNGLMFEWGGISVDPNREVAIANPMALPFSKILIPRGPNMEQPKDAKT 635
QY 375 ---DSGI-----YFIPLNVCYDMAVADQETSMQVNTSNTKLPKPKDMIGRIDAID 425
DB 636 GTESGIQPOYGVPTVTLN-----PFLS-----PFGLPCKOPAMGVIYSLAD 676
QY 426 ISTGRITMSVERAANYS-----PV-----LSTGGVLENGGT--DRYPRAL 465
DB 677 LKTNEVVMKKRIQTPDQSMFPMPVPVFNMGMPMLGPISTAGNVLEFIATADNYLRAV 736
QY 466 SOETGETLMOQLATVYASGA--ISYEDGMQYAI-AGGVSYGS 508
DB 737 NMSGKELWGRLP--AGQATPMTEYVNGKQIVVISAAGHSFST 780

```

RESULT 11
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

RESULT 12
AG0523
glucose dehydrogenase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (S)
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*

[illegible]

RESULT 13

Query Match	12.08;	Score 329;	DB 1;	Length 801;
Best Local Similarity	21.98;	Pred. No. 1e-15;		
Matches 152;	Conservative 81;	Mismatches 197;	Indels 264;	Gaps 33

Db 695 IGITRDSLPFLQLPAYKIGVPLGGSGISTAGNMFLGATQDNYLRAEFTVINGKMLEAR 754

QY 478 LATVASGQA--ISYEVDNQYVAI--AGGVSYSN 508

Db 755 LP--AGGQATPMTYELNKKQYVIMAGHSSEFT 786

QPKEX

Query Match	10.7%	Score 294.5;	DB 1;	Length 808;
Best Local Similarity	20.7%	Pred. No. 3.1e-13;		
Matches 145; Conservative	82;	Mismatches 234;	Indels 241;	Gaps 27

[illegible]


```

QY 265 TFEAMVTNVDPSTFEMGLOSINPNAATGERVLTGPKCTGIMQFDETG-EFILMAR 323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 PSQPSIVDTOKDGLVPAIVA-----PRTGDI FVLDRTKGEIVPAP 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 DT-----NYQNMIESIDENGIVTVNEDALIKELDY-----EYD 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 518 ETEVPQGAAPGDHTSTETQMSQ---LTLRPKNPLNDSDINGTIFDQMFCSYFHTLRYE 574
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 -----VCPFLGGRDMPSAALNPDGIFYIFPLNNVCYDMAVDO----- 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 575 GPETPSLKGSLFPGDLQMFEMGGLAVDPQROVAFANPDISLPEVSQLVPRGPNLMP 634
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -----ETSDVYNTSNVTKLPKGM-----GRIDAIDISTGRTL 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 ENAKGTGETGLQHNNGIPYAVNLHPFLDPLVLPFGIMPKCTPPWGVAGIDELKTNV 694
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 433 MSVERAANYS-----PV-----LSTGGVLE-NGGIDRYFRALSOETG 470
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 WQHRNGTLDMSWSSSLPIPLPIKIGVPSLGLPLSTAGNLGFLTASMDYYIRAVNLTTG 754
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 471 ETLMQTRLATVASGAISTEVDGMQYVAIAGGVSYGSLNS 512
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 755 KVLMDRLPAGAOATPTITAINCKOYI-----VTYAGSHNS 790
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

```

F83360
glucose dehydrogenase PA2290 [Imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83360
R:Stover, C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
  .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: F83360
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <STO>
A:Cross-references: GB:AE004654; GB:AE004091; MID:g9948311; PIDN:AA05678.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```

```

Query Match 10.28; Score 279; DB 2; Length 803;
Best Local Similarity 20.7%; Pred. No. 4.1e-12;
Matches 152; Conservative 79; Mismatches 226; Indels 276; Gaps 32;

```

```

QY 6 LLM-----ASAGALALLAPAFQVPTVDLLANP----- 36
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 LLMLEFRRPLADGPAPLGTALCAVAVLAGAAVGSQ-FTNPGLVGRIDRDSGMTST 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 37 ----PAGEMISYQONENYRHSPLTOITTEENVGQLQVWARGMPGKV-----Q 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 APAMDGDMQAVGRTEFGDRYSPLKQITPANVGQLEBAMR--IRTGDLPTADDPLLELTNE 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 VTPLIHDGVMTLANPDVIOAIDAKTGDLTWEHRQLPNATLNSFGEPT-RGM----- 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 NTPLVNGLVACTASHSVKSLALDPTGAETWRFDPOIOSPVGKFAHMTORGVSYTDEE 279
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 -----AIYANGVIVA-----GSTCQ----- 149
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 QYARSDVGAAPPALSEAKKAAVASCPRRFLPTADARLIAINADNGKVCEDFGVKGAVDL 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 150 -----YSDFGCF-----VSQH---DSATGEE--LMKNYFIPRAGEGDETW 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 TAGIGFTPGGYSTSPAAVTNLTIVLIGSHVTDNESTNEPSGVIRAPDY---HDGKLW 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 GND-----YEARMTGAMGOITVDPTNLVHYGSTAVGPASETORG---TP 228
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 396 NMDSGNPDETEPLADGKFTYTRNSPMMMSLASVDEKLGQV---LPLDNQPDQMGKRT 452
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 GGTLCGTNTRFAVRDGTGEIYWRHQTLPDNNMDQCFEMAVTVNDVDPSTFEMGLOSIN 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 453 GAEKPSAGL-VALDINTGKLRNRYOFTFHDLDMD-----VGSQPTL-----LD 495
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 PNAATGERVLTGVPCKTGTMTMOFDETG----- 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 LKTAQGVKPALI-APTQGSILYVLDLRDGTPIVPIREVPAPQGAVEGDHPTAPQARSDLN 554
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 318 -----EFLMA-----RDYNQNMIESIDENGIVTVNEDALIKELDY 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 LNRPLTERDMMGSSPEFQMLCRIOFRLRYBGQYTPSEQSL----- 598
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 VCPFLGGRDMPSAALNPDGIFYIFPLNNVCYDMAVDOEFT---SMDVYNTSNVK--- 410
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 IYGNVGVFNWGSVDVPRQLFTSPNYMAFVSQMPDRDKVPSGSKREGETSGVQNTG 658
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 -----LPPKDMIGRIDAIDISTGRTLSVERAA-NYSPV----- 445
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 659 APYAVIMHPFMSPIGLPCQAPSWGDVAGIDLTITAKVWQHKNGTISRDMTPVPIGLTVGP 718
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 ----LSTGGVLEFNGT-DRYFRALSOETGETLMQTRLATVASGAO--ISYE-VDGMQY 496
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 719 SMGGSITTAGVAFLSGTLDOYLRAVDVADKQIMQARLP--AGQATPMYSYTKDGROY 776
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 497 VAI-AGGVSYGS 508
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 777 VLIYAGHGSFGT 789
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 24, 2002, 10:21:10
 Job time: 348 sec

Fri May 24 11:27:33 2002

wallick-934-135.ppt.rpr

```
Query Match      17.8%; Score 488; DB 1; Length 739;
```

Best Local Similarity 26.0%; Pred. No. 8.9e-29;
Matches 152; Conservative 76; Mismatches 246; Indels 110; Gaps 14;

```

OY 10 SAGALLAALAPAFQVPTVDE-----LLANPAGCEWISYQONENYHSLPQITTT 61
DB 17 TAGTICAAALISGYATMASADGOGATGEAIHADDHGNMWTYGRYSEGRYSPLOQINR 76
OY 62 ENVGOLQIWMARGOPGVQYV-TPLIHDGVWYLANPGDVIAQIDAKTGDILWEHRQLP- 119
DB 77 SNVGNLKLAWYLDLDTNRGOGCTPLVDGVYATTTNMSMKAVDAATGKLMSYDPRVPG 136
OY 120 NATILNSGEPTRGMAI----- 136
DB 137 NIADKCCDVTYNRGAAYNGKVFEGTEPDRLLALDAKTGLVMSVNTIPPEALGKORSY 196
OY 137 -----VANGVIAGSTCOYSPGCC--FVSGHDSATGEELMRYNITPRAGEBDETWGN 187
DB 197 TYDGAAPRIAKGRVILGN--GSGSEFGARGFYATDAETGKVDWRFETAPRKNPDHTASD 254
OY 188 D-----YEARWMTGA-----WQITYPVTNLVHYGSTAVGPASETQRTPGGT 231
DB 255 SVLNKRAYQWTSPTGANTRGGGGTWDSIVYPVADLYLVGNGSPNNYRSEBKGCD 314
OY 232 LYGNTEFAVAPRTGELVYHQTLPNDNDQECTFEEMVTNVDPSTMEGLQISINPA 291
DB 315 NLFTGSIVALKPTEGEVWHFOETPMQDFTSVQOIMTLDPINGET----- 362
OY 292 ATGERRVLTGVPCKTGMQFQDAETGEPLMARPTNQNIESID-ENGIVTYNEDAILKE 350
DB 363 -----RHVIYHAP-KNGEFTIILDAKTEFTISGKNYVYVNMASGLDPTGPIYNPALYTL 417
OY 351 LDVEYDVCPTEFLGDRWPSALNPDGSIYFPLNNVCYDMAVDOEFT-SMDVYNTS--- 406
DB 418 TGEWGIQIPDGLGHNFAAMAFSPKGLVYIPAOQVPLTYNQGGFTPHRPSNMLGIDM 477
OY 407 NVTKLPBG-----KDMIGRIDALDISTGRTLSVEPAANISPVLSGGGVLENGGT 458
DB 478 NKAGIDSPBAKQAFVKDKLGMVAMPDQKQAEAMRVHKKFPWNGGILATGDLDFQCLA 537
OY 459 DRYFRALSOETGETLQWTRLATVASGOAISYEVDGMQVYAIAGG 502
DB 538 NGEFHAYDATNGSDLFHFAADSGIAPRYTYLIANKQYVAIVEVG 581

RESULT 2
DHET ACEPO STANDARD; PRT; 738 AA.
AC P28036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADH4.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OX Acetobacter.
NC NCB1_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NB11028;
RX MEDLINE=91159482; PubMed=2001402;
RA Tanaka T., Fukaya M., Takemura H., Tayama K., Okumura H.,
RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
RT "Cloning and sequencing of the gene cluster encoding two subunits of
RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.

```

```

CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: D00635; BA00528.1; -
DB PIR: S14270; S14270.
DB HSP: Q924J7; 1PLG.
DB InterPro: IPR001479; Bac_POO.
DB InterPro: IPR002372; Bac_POO.
DB InterPro: IPR000345; CytC_heme_bind.
DB Pfam: PF01011; Bacterial_POO_6.
DB PROSITE: PS00363; BACTERIAL_POO_1; 1.
DB PROSITE: PS00364; BACTERIAL_POO_2; 1.
DB PROSITE: PS00190; CYTOCHROME_C_1.
DB Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
FT CHAIN 1 35
FT BINDING 36 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BINDING 650 653 HEME (COVALENT) (BY SIMILARITY).
FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92A96 CMC64;

```

Query Match 17.6%; Score 483.5; DB 1; Length 738;
Best Local Similarity 25.8%; Pred. No. 1.9e-28;
Matches 152; Conservative 76; Mismatches 239; Indels 123; Gaps 14;

```

OY 10 SAGALLAALAPAFQVPTVDE-----LLANPAGCEWISYQONENYHSLPQITTT 61
DB 17 TAGTICAAALISGYATMASADGOGATGEAIHADDHGNMWTYGRYSDORYSPLOQINR 76
OY 62 ENVGOLQIWMARGOPGVQYV-TPLIHDGVWYLANPGDVIAQIDAKTGDILWEHRQLP- 119
DB 77 SNVGNLKLAWYLDLDTNRGOGCTPLVDGVYATTTNMSMKAVDAATGKLMSYDPRVPG 136
OY 120 NATILNSGEPTRGMAI----- 136
DB 137 NIADKCCDVTYNRGAAYNGKVFEGTEPDRLLALDAKTGLVMSVNTIPPEALGKORSY 196
OY 137 -----VANGVIAGSTCOYSPGCC--FVSGHDSATGEELMRYNITPRAGEBDE----- 182
DB 197 TYDGAAPRIAKGRVILGN--GSGSEFGARGFYATDAETGKVDWRFETAPRKNPDHTASD 254
OY 183 -----ETWGNDEARMWMTGA-----WQITYPVTNLVHYGSTAVGPASETQRTPGGT 291
DB 255 SVLNKRAYQWTS-----PTGAMTRGOGGTWDSIVYPVADLYLVGNGSPNNYRSEBKGCD 314
OY 226 GTPPGTILGINTRRVAPRTGELVYHQTLPNDNDQECTFEEMVTNVDPSTMEGLQ 285
DB 308 SEGKDNLELSIYALKPTEGEVWHFOETPMQDFTSVQOIMTLDPINGET----- 357
OY 286 SINPAATGGERVLTGVPCKTGMQFQDAETGEPLMARPTNQNIESID-ENGIVTYNEDAILKE 344
DB 358 -----NGEHRHVIYHAPKNGEFTIILDAKTEFTISGKNYVYVNMASGLDPTGPIYNPALYTL 410
OY 345 DALIKELDEYDVCPTEFLGDRWPSALNPDGSIYFPLNNVCYDMAVDOEFT-SMDVY 403
DB 411 DALYTLGSKWYVIGIPDGLGHNFAAMAFSPKGLVYIPAOQVPLTYNQGGFTPHRPSNML 470
OY 404 NTS---NVTKLPBG-----KDMIGRIDALDISTGRTLSVEPAANISPVLSGGGV 452
DB 471 NLGIDMKVGIPIPSPEAKQAFVKDKLGMVAMPDQKQAEAMRVHKKFPWNGGILATGDL 530
OY 453 LFNGGTDRFRALSOETGETLQWTRLATVASGOAISYEVDGMQVYAIAGG 502

```



```

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86375; BAA19753.1; -.
DR HSSP; Q9ZAJ7; 1F1G.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR InterPro: IPR000345; Cyt_c_heme_bind.
DR Pfam; PF01011; Bacterial_POQ_6.
DR PROSITE; PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE; PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 34
FT CHAIN 35 757
FT MOD_RES 35 35 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT ACT_SITE 342 342 PYRROLIDONE CARBOXYLIC ACID.
FT BINDING 653 653 BASE (POTENTIAL).
FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 757 AA; 82966 MW; 39B9F90E3B947581 CRC64;

Query Match 17.3%; Score 473.5; DB 1; Length 757;
Best Local Similarity 25.3%; Pred. No. 1.1e-27;
Matches 155; Conservative 88; Mismatches 233; Indels 137; Gaps 19;

OY 7 LMSAGALILAA--PAFAOV---TPYDELLANPPAGEWISYONENYRHSPLTOIT 61
DB 16 LLSCAALAFSAAPVAFVAFQEDGTATITSSDNGHP--DDWLSTGRSYSEORHSPLOQINT 74
OY 62 ENVGQLOLVWARGMOPGKQV--TPLIHDGVYLANPGVIAIDAKDGLLWEHRROP- 119
DB 75 ENVGKIKLAAHVDLTNRQEGSTPLIVNVMKATITNSKMAALDAATGKLSMDPKVPG 134
OY 120 NATILNSEEPTRGMAI----- 136
DB 135 NIADRCCTVSRGAAYMNGKVFYFGTFRDLIALDAKTGLVMSVYTIIPKPAQLGHQRSY 194
OY 137 -----VANGYIVAGSTQVYSPFC--FVSGHDSATGEELMRVFIIPRAGEED- 182
DB 195 TYDGAAPRIAKGVILGN--GGAEEFARFVSAFDETSKLDWREFYVNPENKPKDGAASD 252
OY 183 -----ETWGNDEYEAAMWTG--AMQITTYDVTNLVHSGTAVYPAPESTGRTPGGT 231
DB 253 DILMSKAVPTMKNAMKMOGGGGTVMDSLYDPTDLVYLGANGSPMYKRFSEKGD 312
OY 232 LYGTVTRAVRPDGEIYWRHOTLPBROWDOECTFEEMAVTNVDVOSTEGLQSLNPNA 291
DB 313 NLEFLSIVAINPDGKYWHFQETPMDEMDYTSVOQINTLDMPY----- 356
OY 292 ATGE--RRVLTGVPCKTGTMTGMOFDAETGEFLWARTNYONNIESIDE--NGIYVNEAAILK 349
DB 357 -NGEMRHVIAHAP--KNGEFTIIDAKTGKFTIGKRYTTENNANGIDPVYTGPNVVPDLMT 414
OY 350 ELDEVEYDVCPTFLGGRDMPASALNPDGSIYFIPLNNV-----CYDMAVDQE 396
DB 415 LTGRKWLGIPELGEAGHNEFAAAYSPKTKLVYIPAOQIPLLYDQGGKFKYHDAWMLGID 474
OY 397 FTSMDVYNTSVTKLPBGKIMIGRIDALIDISTGRTL--WSEYRAAAYS-----PVY 446
DB 475 MKKIGLFDNDNPEHVAAKDF-----LKVYLGKWTYAMDEKMAPAFTINHKGPWNGGL 528

```

```

OY 447 STGGVLENGSTDRYFRALSOETGELMOTRLATVASGAISYVDGMQVAT----- 499
DB 529 ATAGNVIFQIANGELAFHAYDATNGNDLYSPFASAIAPVYITANGKQYVAVENGMSI 588
OY 500 ---AGGVSYSYG 509
DB 569 YPFLYGVARTSG 601

```

```

RESULT 5
XOXF_PARDE STANDARD; PRT; 600 AA.
ID XOXF_PARDE
AC P29968;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
GN XOXF.
OS Paracoccus denitrificans.
OC Bacteria, Proteobacteria, alpha subdivision, Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RA Harms N.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 380-600 FROM N.A.
RC STRAIN=PD 1235;
RX MEDLINE=92041583; PubMed=1657873;
RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
RA Stouthamer A.H.;
RT "Isolation, sequencing, and mutagenesis of the gene encoding
RT cytochrome c531 of Paracoccus denitrificans and characterization of
RT the mutant strain."
RL J. Bacteriol. 173:6971-6979(1991).
CC -1- COFACTOR: POQ (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34346; AAC4455.1; -.
DR EMBL; M75583; AAA25574.1; -.
DR PIR; A41378; A41378.
DR HSSP; P38539; AAAH.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam; PF01011; Bacterial_POQ_7.
KW Oxidoreductase; POQ; Signal.
FT SIGNAL 1 21
FT CHAIN 22 600
FT ACT_SITE 318 318 PUTATIVE DEHYDROGENASE XOXF.
FT ACT_SITE 318 318 BASE (POTENTIAL).
SQ SEQUENCE 600 AA; 65159 MW; DCA096F1BC5A3ACE CRC64;

```

```

Query Match 16.1%; Score 441.5; DB 1; Length 600;
Best Local Similarity 25.9%; Pred. No. 2e-25;
Matches 157; Conservative 84; Mismatches 220; Indels 145; Gaps 25;

OY 10 SAGALILAAFAFAQVTPYDELLANP-----PAGEWISYONENYRHSPLTOITTE 62
DB 6 NGACIALLMGSTAA-----LANQGRGRQAPQAWAIDMGDVAANRYSTLQIND 56
OY 63 NVGQLOLVW--ARQMOPGKQVTPPLIHDGVYLANP--GVIOAIDAK--TGDILWEHR--RQ 117
DB 57 NVKDLRAVAFPSFSGVLNGH--EGSPVIVGDMYVHTFPKRVFALDLNDGKILMRYEPOQ 115

```

```

RL  Nat. Struct. Biol. 1:102-105(1994) .
RN  [4]
RX  X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RA  MEDLINE=95253818; Pubmed=7735834;
RA  Ghosh M., Anthony C., Hailos K., Goodwin M.G., Blake C.;
RT  "The refined structure of the quinoprotein methanol dehydrogenase
RL  from Methylobacterium extorquens at 1.94 Å."
RL  Structure 3:177-187(1995) .
CC  -|- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC  reduced acceptor.
CC  -|- COFACTOR: POO.
CC  -|- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC  OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC  -|- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC  MEMBRANE.
CC  -|- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
CC  EMBL: M31108; AAA25380.1; -.
CC  PIR: S07908; S07908.
CC  PIR: J00706; J00706.
CC  DR  HSSP: P38539; AAAH.
CC  DR  InterPro: IPR001479; Bac_POO.
CC  DR  InterPro: IPR002372; Bac_POO_repeat.
CC  DR  Pfam: PF01011; Bacterial_POO_7.
CC  DR  PROSITE: PS00363; BACTERIAL_POO_1; 1.
CC  DR  PROSITE: PS00364; BACTERIAL_POO_2; 1.
CC  KW  Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
CC  FT  SIGNAL 1 27
CC  FT  CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
CC  FT  DISULFID 130 131
CC  FT  DISULFID 413 442
CC  FT  ACT SITE 330 330
CC  FT  ACT SITE 626 AA; 68434 MW; 64988DOAFD2AD34C CRC64;
CC  SQ  SEQUENCE

```

```

Db 345 -----KDGKARKLLTHPRNGIYVTLDRDGLVSAANKLDDT--VNVFKSVTLKGTQPR 397
Qy 344 EDAILKELD-VEYDVCPTFLGRDWPMSALNPDGSIYEFLNNVCYMAAIVDEFTSMQV 402
Db 398 DEYGTBMDHLAKDICPSAMGYHNOGHDSDPKRELFEMGINHICMDEPFMLPYRAGOF 457
Qy 403 YNTSVNTVLPYCK-----DMGRIDAIDISTGRTLMSVRAAANYSPVLSTGGVLFNG 456
Db 458 FVGATLNNYPPKGRDROVYEGIQIKAYNAITGDKWEKKEKRRFVAVGCMATAGDLVFTG 517
Qy 457 GTRDFRALSGETGLMOTRLATVASGAISYEVDMQVATAGGVSYSG 509
Db 518 TLGYLAKRSDTDLLMKRIFPSGALGYPTTHKGTQYVAI-----YGVG 565

RESULT 7
DHML_PARDE STANDARD; PRT; 631 AA.
ID DHML_PARDE
AC P1293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
  alpha subunit) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RA MEDLINE-8707969; PubMed-3114231.
RX Harns N., de Vries G.E., Maurer K., Hoogendijk J., Scoutamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
  structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
  reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
  ON METHANOL (IN P.DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
  TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17339; AAA88366.1; -
DR HSSP; P38539; AAAH.
DR InterPro; IPR001479; Bac_POO.
DR InterPro; IPR002372; Bac_POO_repeat.
DR Pfam; PF01011; Bacterial_POO_7.
DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT CHAIN 1 32
FT DISULFID 135 631 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT_SITE 335 335 BASE (POTENTIAL).
SEQUENCE 631 AA; 69799 MW; 0934DC93FPC5730B CMC64;

```

Query Match 15.8%; Score 433; DB 1; Length 631;
 Best Local Similarity 25.3%; Pred. No. 9,1e-25;
 Matches 152; Conservative 91; Mismatches 220; Indels 138; Gaps 25;

```

Qy 4 TSLMASAGALALL-AAFAFAQVTPVDEL--LANPAGEMISYGNQENYRHSPLTOT 60
Db 12 SSLMAVAVMGLAVLTTPATA-----NDQVELAKDPA-NWVTFGQDYNAQNISENTDIN 65
Qy 61 TENNGQDLVAV-ARGMQPKGVOTPLIHGVWYLANP-GDVIOAIDA-KTGDLIWEHR- 115
Db 66 KENKQDLPAWFSFSTGLHGH-EGTPLVVDKRMFIHTPEPNTFALDNEPKILMNKP 124
Qy 116 RQLPNIATLNSFGEPTRGNA-----YANGVT 142
Db 125 KQNTARTVACCDVNNGLAVYWGDDQVAPLIRFOTLDGHIYAMALETETRMIMENDI 184
Qy 143 VAGSTCOYSEF-----GC-----FVSGHDSATGEELMRYTF----- 174
Db 185 KVGSTLTIAVYIKDLVYSGSGAELGVGYATYAVDKSGEKMRAPATGPDELLAED 244
Qy 175 -----PRAGE--GDETWGNDYFARMGTG--KQITDPTNLVHGSTAVGPASTQ 224
Db 245 FNAPNPHYSGKNGLETWEG--AKKIGGNNWYAYDPEVDFYVSGNPAPWNETIN 301
Qy 225 RGPFGTLGTNTFAVRDPTGEIWRHQTLPDNDQECTFEMVAVTVDOPTSTEMGL 284
Db 302 R--PGDNKW-TMAIWGREATTGEAKFAVQKTPHDEM-----YAGVNMALSEQEDK 350
Qy 285 QSINPMAATGERRVLVGVPCCKTGMQPDATGEFLKARDNTYONMIESIDE--NGIYTV 342
Db 351 Q-----GQNRKLLTHPRNGIYVTLDRDGLISA-----DKMDPTVNWKEY 393
Qy 343 NEDAILKELDVEY-----DVCPTEFLGRDWPMSALNPDGSIYEFLNNVCYDMAV 393
Db 394 QLDTGTPVRDPEFRIMDKARDICPSAMGYHNOGHDSDYDERKVFALNHICMDEPF 453
Qy 394 DEFTSMQVYNTVTKLPCKGDM-----TRIDAIDISTGRTLMSVRAAANYSPVLST 448
Db 454 MLYRAGOFFGATLTPKPGPATAERAGAOIKAYALISGEMKEMEREFVWGTMAT 513
Qy 449 GGGVLENGTDRYFRALSGETGLMOTRLATVASGAISYEVDMQVATAGGVSYGS 508
Db 514 AGGLFEYVTLGDTLAKRSDTDLLMKRFLPSVIGHPMYTKHGRQYVAI-----MYGV 568
Qy 509 G 509
Db 569 G 569

RESULT 8
DHML_METOR STANDARD; PRT; 626 AA.
ID DHML_METOR
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
  alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RA STRAIN-ATCC 27866 / DSM 760 / NCIB 11278;
RX MEDLINE-89008094; PubMed-2459109;
RT "Nucleotide sequence and transcriptional start site of the
  Methylobacterium organophilum XX methanol dehydrogenase structural
  gene.";
RL J. Bacteriol. 170:4739-4747(1988).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
  reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
  OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.

```


	Query Marcov	15.7%;	Score 430.5;	DB 1;	Length 626;	
	Best Local Similarity	24.2%;	Pred. No. 1.4e-24;			
	Matches 148;	Conservative 91;	Mismatches 208;	Indels 165;	Gaps	23
Qy	10 SAGALALLA -AP-AFAOVTVYTBELLANPAGE-WTSTGQNOENRHRHPLQITTEWGO	66				
Dd	7 SVSALAMALLPALISSAYAAANKVELSKSDNNVMPKNDSSNNYSBELKOVAKSNVQ	66				
Qy	67 LIDLW--ARGAQPGVOYTPLHDGVNY-----LANPDV-----	99				
Dd	67 LRPAVFESTGLINH-BGARLYVDGKKYVTSPNNTPALDDDDPQHILMDPKPNDA	125				
Qy	100 -----IQAIDAKTGDLIHERRRLPIATLN	125				
Dd	126 RAVACODDIYNRGIAWPGDGKTPLALIKTQLDHRHVALLNAETGETVWK----	173				
Qy	126 SGCEPRGGAIIANYAVIYAAGSTCQSPF-----GC-----FYSGHDSATGEEL	168				
Dd	174 -----VENSDIKVGSTLIITAIPYVKOKYIIIGSSAEIGVAGYLTAIVDTGGVY	222				
Qy	169 WRNYFI-----PRAGEE--GDETWGNDYEARMWTGA---WGQIYYDPY	207				
Dd	223 WRRTAIGPRKDLLADDFNVKNANHAGQGIGLGTATMBED--AWKIGGGTNMGVATADPET	279				
Qy	208 NLVHYGSTAVGASSETORGTGGLTICGNTFEVARPTGELIVRHOTLPRLDNDOCTPE	267				
Dd	280 NLIYFTGPNAPAWNEMFR--PEDNKW-MTLTFGDADLTGAKEICYOKTPREDEWDAGVAV	336				
Qy	268 MATVINDVOPTSEMEBLSINFNATGERRYLVGPCKTGTMQFDAETGEELMAR--D	324				
Dd	337 MM-----PSBGKD-----KDGTRKLILHPRRNGIYVLTLPDTCALVSARKLDD	380				
Qy	325 TYNQNMIESIDENGITYVEDAILKELD-VKYDCOPFLEGRBMPSALNPBGSIYFIL	383				
Dd	381 T--VNVEKTIQDLTGGPVADPREYGTGRMHLLAKDVCPSSAMGYHNHGHSYTPKRELFFMGI	438				
Qy	384 NNVCYDMAAVDDEFTSMDYNTNSVTKLPBGR-----DMIGRIDALISTGRTLVSVER	437				
Dd	439 NHICMEDPERMLPYRAGOGFVGATLNMVPGRGCRONEGEGOLKAVALNATSYKWEKXE	498				
Qy	438 AAANYSPVLSTGGCVLFNGGTDRIFRALISOETGELIMOTRLATLYAASQALISYEVDKOXY	497				
Dd	499 RFVAWGTGLTAGIVLEYFGTGLDYLKARDSDTDGLMKFLKPSAIGIPMTYTHKGTOY	558				
Qy	498 AIAAGGVSYSG 509					

```

Db 559 AT-----YGVWG 565

|||
|||
|||

RESULT 9
EXAA_PSEAE STANDARD; PRT; 623 AA.
AC 09Z4J7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
GN EXAA OR PA1982.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OC NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=99041560; PubMed=9826187;
RA Dielhl A., Wiltzingerode F., Goerlsch H.;
RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
RT homodimer: sequence of the gene and deduced structural properties of
RT the enzyme.";
RL Eur. J. Biochem. 257:409-419(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiser J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=ATCC 17933;
RX MEDLINE=99173751; PubMed=10075429;
RA Schobert M., Goerlsch H.;
RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
RT dehydrogenase.";
RL Microbiology 145:471-481(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
RC STRAIN=ATCC 17933;
RX MEDLINE=20202376; PubMed=10736230;
RA Keitel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
RA Goerlsch H.;
RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
RT Pseudomonas aeruginosa: basis of substrate specificity.";
RL J. Mol. Biol. 297:961-974(2000).
RN [5]
RP FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
RN ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
RN [6]
RP COFACTOR: POQ AND CALCIUM.
RN [7]
RP SUBUNIT: HOMODIMER.
RN [8]
RP SUBCELLULAR LOCATION: Periplasmic.
RN [9]
RP SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
RN [10]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isd-sib.ch/announce/
RN or send an email to license@sib-sib.ch).
```

DR EMBL: AJ009858; CA00896.1; -
 DR EMBL: AF004624; AAC05370.1; -
 DR EMBL: AF068264; AAC79657.1; -
 DR PDB: 1F1G; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 DR Oxidoreductase; PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 RM Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623 QUTNPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SO SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;

Query Match 15.2%; Score 417; DB 1; Length 623;
 Best Local Similarity 23.1%; Pred.No.1.4e-23;
 Matches 140; Conservative 95; Mismatches 247; Indels 124; Gaps 20;

QY 3 PSLIMASAGALAL--LAPAFAYTPTVDLLAN--PPAGEMISYGCNENRHSPLT 57
 DB 9 PACLRPSLHCLAFVALGASAGALAKDVTWEDIANDKTTGDVLYQNGHQAQWSPK 68
 QY 58 QITTEWGOQLVWARGM--QPGVQVTPLIHGVMTLANPGDVIAIDAKTGLIME 113
 DB 69 QVANANVFKLTPMAYSYSGFDEKORQ-ESQAIYSDGVIVTASYSLPALDAKTKRLMT 127
 QY 114 HRRLP-----NATLNS----- 126
 DB 128 YNRRLDDLRPCDDVYVNRRAIYGVFGTLDASVVALNKNKGAVVKKKFAHDGAGYT 187
 QY 127 -PEEPRGAIYANG--VIVAGSTCOYSPGCF--VSGHDSATGELMRYNFIPT- 176
 DB 188 MTCAPF---IYDKGTGKVLIIHGSS--GDEFVYVGRLPARDPOTGSEIMRPFVEGHM 241
 QY 177 -----AGEBDEMTGNDEYA-----RMATGA--WGQITYPVTNLVHYGSTAV 217
 DB 242 GRUNGKSTVYTDVAPSPDPDRNSPTGKVESWHSHGAGPQASAFPAETWTIIVGAGNP 301
 QY 218 GPASRTGRTGG--TLYGNTFRAPVPDTEIYWRHQTLPRNMDCECFEMAVT 271
 DB 302 GPWNTWARTAKGNPHDYSLY-TSGQYGVDPSSGEVYKVFQHPNDAMPDSGNNELVLF 360
 QY 272 NVDVQPTSTEMEGLSINN--AATGERVLTGVCKTGTW--QFDAETGEFLWARD 324
 DB 361 DYKAKDGKIYATVAHADRNGEFFVYVDRSNGKIQNAFPFVNITWASHIDLKTG---RP 415
 QY 325 TNYQNMIESIDENGIVTVNEDAILKELDEYDVCPFLGGRDMSALANDSGITFPLN 384
 DB 416 VEREGORPPLPEPG-----QKHGKAVEVSPPLGGKNNMPPMAYSDDTGLFYVPAN 465
 QY 385 NVCIYMAVADQFTSMQVYNTSVTKLPKGMIGRIDAIDISTGRTLSVERAANYSF 444
 DB 466 HKMEDYMTVEEVSYTKGSAYLGMGRIRKMYDDHVSIRANDPVSGKVVWHEKHELELMAG 525
 QY 445 VLSGGVLENGGTRDYFRALISOETGETLMOITRLATYASQAISYEVDGMOYVALAGGV 504
 DB 526 VLAIRAGNLVFTGTGDTFKAFDAKSGKELMKFQTGSIYSPITWEDGBOYLGVT--V 582
 QY 505 SYGSGL 510
 DB 583 GYGAV 588

RESULT 10
 DHM1_METME STANDARD; PRT; 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Methanol dehydrogenase subunit 1 (Ec 1.1.99.8) (MDH large alpha
 DE subunit) (MEDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 CC Methylophilus.
 OX NCBI_Taxid=17;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=94059969; PubMed=8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RA Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RT methanol dehydrogenase.";
 RL Biochemistry 32:12935-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93054513; PubMed=1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution.";
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 DR PDB: 4AAH; 08-DEC-96.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
 KW Calcium.
 FT DISULFID 103 104
 FT DISULFID 379 408
 FT ACT_SITE 297 297 BASE (POTENTIAL).
 FT HELIX 2 9
 FT TURN 11 12
 FT STRAND 14 14
 FT TURN 17 18
 FT STRAND 21 22
 FT TURN 26 27
 FT STRAND 34 36
 FT TURN 37 39
 FT STRAND 41 47
 FT TURN 59 61
 FT STRAND 62 63
 FT TURN 64 68
 FT STRAND 71 73
 FT TURN 75 79
 FT STRAND 80 81
 FT TURN 83 84
 FT STRAND 86 90
 FT TURN 96 101
 FT TURN 103 104
 FT STRAND 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT TURN 146 148
 FT HELIX 149 149
 FT TURN 151 151
 FT STRAND 157 159
 FT TURN 160 161

RT and regulation of its expression.";
 RL J. Bacteriol. 175:568-571(1993).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110.
 RX MEDLINE-94261430; PubMed-8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [5]
 RP TOPOLOGY.
 RX MEDLINE-93286127; PubMed-8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase of Escherichia
 coli modelled on that of methanol dehydrogenase from Methylobacterium
 extorquens.";
 RL Biochem. J. 312:679-685(1995).
 [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Corlier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of Escherichia
 coli modelled on that of methanol dehydrogenase from Methylobacterium
 extorquens.";
 RL Biochem. J. 312:679-685(1995).
 [7]
 RP FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; X51333; CAA35706.1; -
 DR EMBL; D12651; BAA02174.1; -
 DR EMBL; D26562; CAB20298.1; -
 DR EMBL; AE000122; AAC73235.1; -
 DR PIR; JY0107; JY0107.
 DR HSSP; P38339; 4AAH.
 DR EcoGene; EGI0369; gcd.
 DR InterPro; IPR001479; Bac_POO.
 DR InterPro; IPR002372; Bac_POO_repeat.
 DR Pfam; PF01011; Bacterial_POO_7.
 DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 11 37 PROBABLE.
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).

FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 141 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TUSADATP -> HIKRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 O -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;
 Query Match 12.7%; Score 349.5; DB 1; Length 796;
 Best Local Similarity 22.8%; Pred. No. 2.1e-18;
 Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;
 QY 6 LLMAS-----ACALALPAFAVQVPTVDLLANPPAGEWISYQNGVRRSRPLTQ 58
 DB 135 LTMAGFNDPQEIINGTILSADATPAEA-ISPVAQD-----DMPAGRNGEQGRSPKQ 185
 QY 59 ITTENAGOLQLVMA-----RGMQPKV--QVTPLIHDCWYLANPGDVIOAIDAKTGD 109
 DB 186 IMADVNHLKEAWVFRTGDVKNPDGELTNEVTPIKVGDTLYLCTAHQRLFDALDAAGK 245
 QY 110 LIMEHRQLPNIATLNSFG-----
 DB 246 EKHMYD---PELKTNESFOHVTGCVSYHEAKETAPEVMADCPRIITLPVNGRLIAI 302
 QY 129 -----EPTRGMAIVANGVIAGS-TCQYS--PFG 154
 DB 303 NAENGRKLCETFAKNGVNLQSNMPTKPGLYEPTSPITTDKTIYAGSVYDNFSTRETS 362
 QY 155 CFVSGHSAAGELNNRYFPRA-----GEGDETCNDYEARMWGMAGQITYPEVT 207
 DB 363 GVIRGEFVNTGELMA--DPCGAKDPNALPSDEHTFTNS-----PNSNAPAYDAKL 413
 QY 208 NLVHGSTAVGPASEFQRTGPGTLYGTNTPFAVPDPDGEIYVWRHQTLPNNMQDECFEE 267
 DB 414 DIVYLPNGVTPPIWNGNTPREERYASSI-LALNATGKLKMSQYVHHDLMDMD--- 468
 QY 268 MMYTNDVOPSTEMESLOSINPNAATGERRVLTGVCPTGTWMOPTAGTEFL----- 320
 DB 469 ----LPAGPT-----LAQITVA--GQKVPVIAPAKTNIFVLDNRNGELVYPAPEKP 515
 QY 321 ----WADTNYQNQNIIESIDNGI-----VYNEDAILKEL--DVEYD----- 356
 DB 516 VPQGAAGKGVYTPQPSLSRPTKDSGADMGATMFDOLVCVWFHQRMYGIEFTPP 575
 QY 357 ----VCPFLGGRDMPSPALNP-----
 DB 576 SEQGLVFPNGNLGMPFEGWGISVDPNREVALINPMLPVSUKLIRPGONPMEQKDAKT 635
 QY 375 --DSGI-----YFPLNNVCYDMMAVDQFTSMQVYNTSNVTKLPFGKDMIGRIDAI 425
 DB 636 GTESGIQPGYGVYGVTLN-----PFLS-----PGLPCKQPAWGYISALD 676
 QY 426 ISTGRTLVSEVERAANYS-----PV-----LSTGGGVYFNGCT--DRYRRL 465
 DB 677 LKINEVWKKRIGTPODSMPFMPVPVPPNMGMPMLGSPISATGNNVLIATATDNYLRAY 736
 QY 466 SOETGETLMQTRLATVASQA--ISYEVDGQGYAI--AGGSVSYGS 508
 DB 737 NMSGKELMQRLLP--AGGQATPMTEYVNGQYVVISAGHGSGTCT 780
 RESULT 12
 DHGA_ACICA STANDARD; PRT; 801 AA.
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor

DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH-A.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 NC NCBL_TaxID=4771.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMD 79.41;
 RA MEDLINE=88289368; PubMed=3399393;
 RT Cleton-Jansen A.-M., Odele G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus";
 RT Nucleic Acids Res. 16:6228-6228(1988).
 RL -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X07235; CAA30222.1; -
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
 Query Match 12.08; Score 329; DB 1; Length 801;
 Best Local Similarity 21.98; Pred. No. 7.2e-17; Indels 264; Gaps 33;
 Matches 152; Conservative 81; Mismatches 197;
 QY 20 PAFQVPTVTEDELANPAGEMISYQONQENYRHSPLTQTTEENVGOLQVWARGQPK 79
 DB 152 PETAQVGVAF-----SMPAYGRTQAGVRSPLKQINDQNVADLKVAWT--LRTGD 202
 QY 80 V-----QVPLIHGVMTLANPGDVIQAIDAKTDLIW-----EH-- 114
 DB 203 IKTNDSEGTNQVTPKIGKNMFCITAHOOIAIDPATGKEKREDPKLTKDSFQHLT 262
 QY 115 -----RRQLP-----NIATINS-----FGE-- 129
 DB 263 CRGWYTDANNTTFRATSLQSKSSQCRPKRVVPVNDGRIVAVNADTKACTDQGNG 322
 QY 130 -----PTRGMAIVANGVIVAGS--TCQYS--PFGCFVSGHDSATGEEL 168
 DB 322 QVNIQEFMPYAVPGGYNTPSGIYTGSVIVVAGSVTDNYSNKEPSG--VIRGYVNTGKEL 381
 QY 165 WNYITPRAGE-----EGDETGNQDYEARMTGAMQOIYIDPVTNLVHGTAIVGPA 222

DB 382 W--VFDTGAADPNAMPGRQETFEVHNS-----PNAMADLADAKLDIV-VYPTGV----- 427
 QY 223 TORGTPGGGLY-----TNTFAVRPDGTGEIWRHOTLRDWMDOCEFEAMVTN 272
 DB 428 ---GTP--DIWGDRTLEKERYANSMLAINSTGKLVNFOPTHHDLDMDVPSQPSLAD 482
 QY 273 VDVQSTEMEGLOSINP--NAATGERRVLTGV-----CKTG--TMMQPAET 316
 DB 483 IKKAGQYTPALIVLTIGNAFVLDLR--NGQPIVPTREKPVQTVKRGPOTKGEFFSKT 540
 QY 317 GEF-----LMA-----RDTYQWMIESTIDENGIVTVNEDAIL 348
 DB 541 QPFSDLNLPADKLDKDKDMWGTMTLDLMCRVSEFRKLVDGILYPPPSBNGTL----- 592
 QY 349 KELDVEDVDCPIFLGGRDPSAALNPD-----GIYFI----- 381
 DB 593 -----VPPGNLVGFEMGMSVNPDRQVAVMNPGLFVSRLLPADPNRAQAKAGT 644
 QY 382 -----PLNNVCYDMAVVDQFTSMQVNTSVTKLPPGKDMIGRIDAIDISTGRTW--- 433
 DB 645 EGVQPMYGVPY-----GVEISAFSLPLGLPCQAPAWGVAGVDLTHEVWKKR 694
 QY 434 -----SVERAANYSPLVSTGGGVLENGT--DRYFRALSOETGETIMQTR 477
 DB 695 IGTIRSLPMLFQLPAVKIKGVPLGGISTAGNVMFVATQDNTLRANVTNGKKLMEAR 754
 QY 478 LATVASGA--ISEYDGMQYVAI--AGGVSYS 508
 DB 755 LP--AGGQATPMYELINGKQYVIMAGHGSFQT 786
 RESULT 13
 DHG_GLUOX
 ID DHG_GLUOX STANDARD; PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 NC NCBL_TaxID=442.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RA "A single amino acid substitution changes the substrate specificity
 of quinoprotein glucose dehydrogenase in Gluconobacter oxydans";
 RA Mol. Gen. Genet. 229:206-212(1991).
 RL [2]
 RN REVISION TO 213.
 RP Goosen N.;
 RA Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

```

CC EMBL: X62710; CAA44594.1; ALT_SEQ.
DR PIR: S17716; OPKEX.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
KW Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
FT SIGNAL 1 33
FT CHAIN 34 808
FT TRANSMEM 35 54
FT TRANSMEM 59 76
FT TRANSMEM 94 108
FT TRANSMEM 123 138
FT ACT SITE 470 470
FT VARIANT 788 788
SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

```

Query Match 10.7%; Score 294.5; DB 1; Length 808;
 Best Local Similarity 20.7%; Pred. No. 2.7e-14;
 Matches 145; Conservative 82; Mismatches 234; Indels 241; Gaps 27;

```

QY 13 ALALLAPAPAPQVTPVDELIA-----NPPAGEWISYQONQENYRHSPLQITTT 61
DB 128 AVLLAFSLFTDPHDSIGELPTQIANASPADPNVPASEWHAQGRQADGRMSPLNQIA 187
QY 62 ENVGOLDLVW-----ARQMGKV--QVTPLIHDSVWYLANPQVYQAIQAKTGDLIW 112
DB 188 TNSNIAKVAHHIHKDKMANSNDSEQINENATPLEFNNTLYMCSLHOKLPAVGAAGVNWKM 247
QY 113 EHRROL--PIATLN---SFGE-PTKGM-----AIYANGVIV----- 143
DB 248 VYDEKLDINPGFHLTCRGVSHFETPANAMDSGNPAPIDCAKDSLIPVNDRLVEVDAD 307
QY 144 AGSRC-----QYSPRCFV----- 157
DB 308 TGKTCSEFGNNGELDLRVNPQPTTGGYEPTSPYITDKLIANSATDNGSVKQASGA 367
QY 158 -SGHDSATGEELW---RNYFIPRAGEGDETWGNDYEARMWTGAMGQITTDPTNIVHY 212
DB 368 TQADVDTGRKRVWFDAISNDPQDLDESHPRVHPHSPMSWYS-----SIDANILVYI 422
QY 213 GSTAVGASSETQGTG---GGLTYGNTNR---AVRPDTEIYWRHQTLLPRNDWDEC 264
DB 423 PMGV-----GTPDQMGDGRTKDSEFFAGIYALNADTKLAMEFYQVYHDLMDLMEI 473
QY 265 TPEMAYTNVDPSTMEGLOSINPAATGERRVLTVGVCCKTGTMQDAEG-ELMAR 323
DB 474 PQQPSLVDTYQKDGTLVPAIYA-----PTKGDIFVLDRTGKEIVAP 517
QY 324 DT-----NYQNMIESIDENGIYTVNEDAILKELDY-----EYD 356
DB 518 EFTVPOGAAPBGDHTSPQMSQ---LTLRPKNPLNDSDIWGTIFDQFCSIFHTLRYE 574
QY 357 -----VCPFLGGRDMPKSAALNPDGSIYFIPLANNVYDMAVDO----- 395
DB 575 GPTTPSLKSLIFPDGLDFEMWGLAVDPQOVAFANPISLPEVSQQLVPPGPNLWPE 634
QY 396 -----EFTSMVYNTSNVTKLPPGKDM-----GRIDAIDISTGRTL 432
DB 635 ENAKGTGSETGLQHNNGIYAVNHLPLDVLTPRGICKPCTPWWGIVAGIDIKTKNV 694
QY 433 WSTERRAANYT-----PV-----LSTGGVLF--NGGTRFYRALSETG 470
DB 695 WQHRNGTLLDSWYGSLLPIPLPIKIGVPSLGLSTAGNGLFLTASMDYITRAYNLTTG 754
QY 471 ETLMTRLATVASGAISYEVDSQYVAIAGGVSYSGSLNS 512
DB 755 KYLMQDRLLPAGQAPITVYALNGKQYI-----VYAGGHNS 790

```

```

RESULT 14
QY1A_ACICA STANDARD; PRT; 809 AA.
AC Q59086;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
GN QY1A.
OS Acinetobacter calcoaceticus.
OC Bacteria: Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.A. AND CHARACTERIZATION.
RC STRAIN=BD413 / ADP1;
RX MEDLINE=95095936; PubMed=8002591;
RA Elsmore D.A., Ornston L.N.;
RT "The pca-pob supraportonic cluster of Acinetobacter calcoaceticus
RT contains quia, the structural gene for quinate-shikimate
RT dehydrogenase."
RL J. Bacteriol. 176:7659-7666(1994).
RN [2]
RP SEQUENCE OF 1-20 FROM N.A.
RC STRAIN=BD413 / ADP1;
RX MEDLINE=96011389; PubMed=7592351;
RA Elsmore D.A., Ornston L.N.;
RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
RT in Acinetobacter calcoaceticus."
RL J. Bacteriol. 177:5971-5978(1995).
CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
CC dehydroquininate + reduced pyrroloquinoline-quinone.
CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
CC -1- COFACTOR: PQQ.
CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY: FIRST STEP. THIS PATHWAY
CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOGLUTARATE
CC PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- INDUCTION: BY PROTOCATECHUATE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: L05770; AAC37161.1; -
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE NEG.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE NEG.
KW Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
FT TRANSMEM 14 34
FT TRANSMEM 41 61
FT TRANSMEM 68 88
FT TRANSMEM 90 110
FT TRANSMEM 127 147
SQ SEQUENCE 809 AA; 88196 MW; 71F67CEBFA62BFCB CRC64;

```

Query Match 9.3%; Score 254.5; DB 1; Length 809;
 Best Local Similarity 21.8%; Pred. No. 2.6e-11;
 Matches 148; Conservative 69; Mismatches 249; Indels 213; Gaps 25;

```

CC -1- COFACTOR: PQQ (b1 SIMILARITY). PATHWAY: FIRST STEP, THIS PATHWAY
CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY: FIRST STEP, THIS PATHWAY
CC -1- ALLONS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
CC PROCOFAEQUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
CC PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on les
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF109471; AAD38453.1; .
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ repeat.
DR Pfam: PF01011; Bacterial_PQQ_6.
DR PROSITE: PS00363; BACTERIAL_PQQ_1, FALSE_NEG.
DR PROSITE: PS00364; BACTERIAL_PQQ_2, FALSE_NEG.
DR Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
KM TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 77 94 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
SQ SEQUENCE 790 AA; 82896 MW; B75F2B52A4F9E6F CRC64;

Query Match 8.6%; Score 235; DB 1; Length 790;
Best Local Similarity 21.3%; Pred. No. 7.3e-10;
Matches 150; Conservative 68; Mismatches 234; Indels 252; Gaps 31.

QY 9 ASAGALALA--AAFAQVTPVYDE-----LIANPPAG--EWSYGOENQYRRHSPTQI 59
DB 110 ATVALAVVAGIGEMFPHPPVAGNAGPGGATVAPGSSVOQNSATAGNTDGGSRRAALDOI 169
QY 60 TTEVNGQQLVWARGMOPGRV-----QVPLIHGVWYLANPGDVIOAIDAKTGD 109
DB 170 NRSN--GRPAAGSPPTTPEGLIANSDDGABDQLTFLVGEKEKVFELCTPHNNLIADASTGK 228
QY 110 LIWEHRRQLPIATLNSFGFPTGMAI-----VAN-----GYVA--GSFCQSP 152
DB 229 QLM--RRE--INATSSWQRCRLGFEADAAALPAPVANSPIAAVTVAOGANCRRRL 283
QY 153 ECCEFSGH-----DSATGEBELRNFTIPRAGEED----- 182
DB 284 FNTNTDGLRIAYADDTGA-----FCQGFSGNSQVLDLKGGLCAARDPFLYLTSPPLYAGT 337
QY 183 -----ETW-----GNDYE-----ARMIGAW 198
DB 338 TVVGGRTRADDNVQITDMGCVYRGSMSPVRSAGLDLPGNMHDROAPAAAGSSVYRSTPNW 397
QY 199 GQITTPDTNLYHNGSTAVGPASTQKGTGPGALLYGCN-----TRFAVRPDTGEI 248
DB 398 APMASDAAMNIVF-----LPLGGESTLDLYGAERTALDHRGASVALADATGAE 446
QY 249 VNRHQTLPEDMNDQCTFEMAVTVNVDOPTSTEMEGQLQSTINPNAATGERRVLTGVCKTGT 308
DB 447 KMYQYQVHNHMD-----FDLPMPQSL-----IDFPMODGSHTPAVAVICTKAGQ 490
QY 309 MMQFADEGEFLW--ARDT-----NY-----QNMESIDENGATYVVEDAILKELVE 354
DB 491 IYVLDRAATGKPLTEVERPVKGSIDIAHEQYAPLOPLSVGMPQIGTKHLTLESMDGATAMD 550
QY 355 YDVC-----PTFLGRDMPASALNPDGSIYF----- 380
DB 551 QMCIKRAIFQOMYBELTYAPGTDVSLSPGSLAGMGGLSTPVDHYVAFADMRIGLWY 610
QY 381 --IPLNN-----YCYDMAVYDOEFLSKDQYVYTSNVTKL-----PGKDMIGRI 421
DB 611 QMIPATRTAAEAAGGEAVNTGMGAAPVLKGPYAVVKNRFLSLALGIPCOAPP-----YGTIL 666

```

```

OY 422 DAIDISTGRTLSVFRRAAN-----VSPV-----LISTGGVLFNGGT-DRY 461
    ||: | | |
DB 667 SAIDIKTRSIAMQVPVGTVDTPGPGIKMLPLPIGMPITLGTLSTOGGLVFIAGTODYY 726
OY 462 FRALSOETGETIMOTRLATVASGOAISY--EVDGMOYVAIAGG 503
    ||: | | |
DB 727 LRAFDSATGKEIMKGRLPVGSOGGPITYVSHKGTGKQYVVISAGG 770

```

Search completed: May 24, 2002, 10:30:56
 Job time: 634 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:19 ; Search time 93.8 Seconds
(without alignments)
944.279 Million cell updates/sec

Title: WALICK-934-135.PEP
Perfect score: 2742
Sequence: 1 MKPYSILMASAGALALIAAP.....GMQYVAIAGGVSYSGILNS 512

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1511	55.1	608	2	093RE9
2	527.5	19.2	698	2	09KH03
3	495	18.1	742	2	053362
4	488	17.8	601	2	09EYH8
5	487	17.8	601	16	092WY9
6	485.5	17.7	708	2	046444
7	482.5	17.6	695	2	09F902
8	470.5	17.2	601	2	09F109
9	466.5	17.0	629	2	09A048
10	462.5	16.9	691	2	09AP95
11	428.5	15.6	633	2	024759
12	420	15.5	623	2	09AGW3
13	419.5	15.3	599	2	09L935
14	392.5	14.3	695	2	0934G0
15	315.5	11.5	573	2	059540
16			790	2	09X2S5

17	290	10.6	785	16	098KF6	098KF6 rhizobium l
18	287.5	10.5	786	2	P95466	P95466 pantoea cilt
19	279	10.2	803	16	091115	091115 pseudomonas
20	263	9.6	777	16	092RB3	092RB3 rhizobium m
21	234.5	8.6	644	2	052551	052551 pseudomonas
22	220	8.0	639	2	P77931	P77931 pseudomonas
23	219	8.0	470	2	030326	030326 acetobacter
24	208.5	7.6	182	2	032699	032699 hyphomicrob
25	203.5	7.4	179	2	032697	032697 hyphomicrob
26	200.5	7.3	180	2	032700	032700 hyphomicrob
27	200.5	7.3	180	2	032700	032700 hyphomicrob
28	200.5	7.3	181	2	032615	032615 hyphomicrob
29	199.5	7.3	181	2	032692	032692 hyphomicrob
30	197.5	7.2	181	2	032703	032703 hyphomicrob
31	196.5	7.2	181	2	032621	032621 hyphomicrob
32	195.5	7.1	182	2	032706	032706 hyphomicrob
33	193.5	7.1	180	2	032701	032612 hyphomicrob
34	191.5	7.0	182	2	032612	032693 hyphomicrob
35	191.5	7.0	184	2	032693	033894 hyphomicrob
36	190.5	6.9	172	2	033894	033882 methylosinu
37	186.5	6.8	172	2	033882	032707 hyphomicrob
38	186.5	6.8	179	2	032707	032694 hyphomicrob
39	185.5	6.8	180	2	032694	032704 hyphomicrob
40	185.5	6.8	180	2	032704	033884 methylosinu
41	185	6.7	171	2	033884	032702 hyphomicrob
42	183.5	6.7	179	2	032702	032695 hyphomicrob
43	182.5	6.7	184	2	032695	0918X8 uncultured
44	181.5	6.6	185	2	0918X8	0918X8 methanotroph
45	176.5	6.4	172	2	093K58	

ALIGNMENTS

RESULT	ID	Query Match	Score	Length	DB 2;	Length	DB 2;
093RE9	093RE9	55.1%	1511	608	2	608	2
AC	093RE9	Best local similarity	50.7%	Pred. No. 1.3e-101			
DT	01-DEC-2001 (TRENBLREL. 19, Created)	Matches 295; Conservative 64; Mismatches 143; Indels 80; Gaps 6;					
DT	01-DEC-2001 (TRENBLREL. 19, Last sequence update)						
DT	01-DEC-2001 (TRENBLREL. 19, Last annotation update)						
DE	ALCOHOL DEHYDROGENASE.						
GN	ADH.						
OS	Pseudoglucobacter saccharoketogenes.						
OC	Bacteri. Pseudoglucobacter.						
OX	NCBI_TaxID=133921;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN-IFO 14464;						
RA	Shibata T., Saito Y.;						
RT	"Alcohol dehydrogenase."						
DR	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AB046580; BAB62258.1;						
SQ	SEQUENCE 608 AA; 65101 MW; 0ACEC97AE11BA570 CRC64;						

QY	10	SAGALALIAAPFAFO-----VTPVDELLANPPAGWISYQONEN 50	
DB	16	STALIASLSGPAFOHDAANAAEFSKAGQSAIENFQVTFADDLGKKNPAMPIIKRNYOG 75	
QY	51	YRHSPLVITITENNQGLQVYARGMGQKQVPTLHIDGYMYLANPQDVIOALDAKTGDU 110	
DB	76	WGYSPLDIDNKDNDGDLQVMSRTMRPSNKGAAIAYNGVIFLGNITVDVIALDGTGSL 135	
QY	111	IMEHRLPNIAT-LNSGPEPTGMA----- 135	
DB	136	IWEYRKRLPSASKFINSIGAAKRSIALFQDKVYFVSMDFVALDAKTGLAMETNRGQS 195	
QY	136	-----IVANGVIVAGSTQYSPFGCFVSGHDSATGEELMRNYFLPRAGEBDEET 184	

```

Db 196 VEEGVANSSGPIVYDGVIASTGQFSGFGYVGTCTAEGSEBELMRNTEFPRGEGEDDT 255
QY 165 WCN-DYEARNMTGAMGQITTPYTNLVHGSTAVGPASETRGTPGGTLVGTNTFAVRP 243
Db 256 WGAAYERKRMWMTGAMGQITTPYTNLVHGSTAVGPASETRGTPGGTLVGTNTFAVRP 315
QY 244 DTGEIWMHQTLPDRNDMDCECTFEMAVTANVDOPSTEMEGLOSINPAATGE-RRLVTV 302
Db 316 KTGEVVMHQTLPDRNDMDCECTFEMAVTANVDOPSTEMEGLOSINPAATGE-RRLVTV 375
QY 303 PCKTGTMOQFAETGEFLMAROTNYQNMIESIDENGLVTVNEDAILKELDEYDVCPTFL 362
Db 376 PCKTGTMOQFAETGEFLMAROTNYQNMIESIDENGLVTVNEDAILKELDEYDVCPTFL 435
QY 363 GGRWPSALNPDSGIFYFLNNVCYDMAVNDDEFTSMOYNTSNVTKLPPGDMIGRID 422
Db 436 GGRWPSALNPDSGIFYFLNNVCYDMAVNDDEFTSMOYNTSNVTKLPPGDMIGRID 495
QY 423 AIDISTGRTLMSVERAANYSPVLSTGGVLENGGTDYRFRALSOETGETLMQTRLATVA 482
Db 496 AIDISTGRTLMSVERAANYSPVLSTGGVLENGGTDYRFRALSOETGETLMQTRLATVA 555
QY 483 SGQAISEVDGMQVVAATAGG-----VSGSGSLN 511
Db 556 SGYTTSTSIDGRTVAVVSGSLGSGPTFGPTTPVDASAGAN 597

```

```

RESULT 2
ID 09KH03 PRELIMINARY: PRT: 698 AA.
AC 09KH03;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update):
DE TERAHYDROFURFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria: Proteobacteria: beta subdivision: Ralstonia group.
OC Ralstonia
OC NOBL_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21122557; PubMed=11222593;
RA Zarrt G., Schrader T., Andreesen J.R.;
RT "Catalytic and Molecular Properties of the Quinolhemoprotein
RT Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo.";
RL J. Bacteriol. 183:1954-1960(2001).
EMBL: AF273737; AAF86335.1; -.
HSSP: Q92437; IFLG.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt_C.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR Pfam: PF00034; cytochrome_c; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180DB1D2FB2 CRC64;

```

```

Query Match 19.28; Score 527.5; DB 2; Length 698;
Best Local Similarity 27.98; Pred. No. 5.4e-30;
Matches 163; Conservative 83; Mismatches 220; Indels 119; Gaps 22;
QY 10 SAGALLAALAAPFA--QVTPYTDDELANPPAG--EWISYQNDENYHSPRLQITTEENVG 65
Db 14 AASVALPAPMAFGANAAARVDGAIRANFAGTNPMSYGLDYAETRFSLKLEQVNAQNV 73
QY 66 QLOIYVWARGMPGK-VQVTPLIHGVVYLANPGDVIOAIADAKGDLIMHRRQLP-NIAR 123
Db 74 NLGLAMSDLESTRGVEATPLVYDGVMTVSAFMSVVAIDARTGKRLMTYDPOVPRDQAV 133
QY 124 LNSGEPPLRGMAI-----VA 138

```

```

Db 134 KCCDDVNRGVALLYQKGFVGAFFDGLVAIDAATGKKWEDDTIVDRSKSYITTCAPRY 193
QY 139 NG-VIAGSTQYSPFCFVSGHDSATGEELMNYFIP-----RAGEGDEMGMD 188
Db 194 NGVILIGGAGLEVGR-VITADAETGKQNMRYVTPGDPARPFENEMAKRAATW--D 250
QY 189 YEAR-WMTG-----AMQIYDPTNLVHGSTAVGPASETRGTPGG-TLVGTNTFAVR 242
Db 251 PSGRVWINGGGGTWNMTAPDELINLMTYTGNGAGSPKRLSPKGDNLVAASV-VALN 309
QY 243 PDTEIWMHQTLPDRNDMDCECTFEMAVTANVDOPSTEMEGLOSINPAATGERVLT 300
Db 310 PDTEIWMHQTLPDRNDMDCECTFEMAVTANVDOPSTEMEGLOSINPAATGERVLT 351
QY 301 GVPCKTGTMOQFAETGEFLMAROTNYQNMIESIDENGLVTVNEDAILKELDEYDVC 358
Db 352 HAP-KNGFEVVIDRTNKFISAKNFVDVWASGDKNGRPHVTPQADTSGKRA----DYV 406
QY 359 PTFILGGRDWPSSALNPDSGIFYFLNNVCYDMAVNDDEFTSMOYNTSNVTKLPPGDMIGRID 407
Db 407 PTFILGGRDWPSSALNPDSGIFYFLNNVCYDMAVNDDEFTSMOYNTSNVTKLPPGDMIGRID 461
QY 408 ---VTKLPPGDMIGRIDAIDISTGRTLMSVERAANYSPVLSTGGVLENGGTDYRFR 463
Db 462 LGMLVNAEPFRSKPMGRLLAMPDLAQAVMHRDHGPMNGGLATAGNLVPGTADGRV 521
QY 464 ALSQETGETLMQTRLATVASGQAISEVDGMQVVAI-GGVSYG 507
Db 522 AYHAATGKLMQAPGTSGVVAAPVYTLIDGRQYVSAVWGGVYG 566

```

```

RESULT 3
ID 053362 PRELIMINARY: PRT: 742 AA.
AC 053362; Q44139;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus (Acetobacter turbidans).
OC Bacteria: Proteobacteria: alpha subdivision: Acetobacteraceae;
OC Acetobacter.
OC NOBL_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94042848; PubMed=8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus.";
RL J. Bacteriol. 175:6857-6866(1993).
EMBL: D13893; BAA40252.1; -.
HSSP: Q92437; IFLG.
DR InterPro: IPR001479; Bac_PQO.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR PROSITE: PS00363; BACTERIAL_PQO; 1.
DR PROSITE: PS00364; BACTERIAL_PQO; 2; 1.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

```

```

Query Match 18.18; Score 495; DB 2; Length 742;
Best Local Similarity 27.08; Pred. No. 1.4e-27;
Matches 164; Conservative 82; Mismatches 221; Indels 140; Gaps 22;
QY 11 AGALLAALAAPFAQVTPYTDDEL--ANPPAGEWISYQNDENYHSPRLQITTEENVG 68
Db 23 AALPVAIVPARADGGGTGATIIHADHPENMLSYGRYSQYSEORVPLDQINRSNGDLK 82
QY 69 LVWARGMPGKVVQ-VTPLIHGVVYLANPGDVIOAIADAKGDLIMHRRQLP-NIAR 122
Db 83 LAWYITLDTNRGQATPLVYDGVMTVSAFMSVVAIDARTGKRLMTYDPOVPRDQAV 142

```

```

QY 123 --TLN-----SRGE-----PTRGMAI 136
DB 143 CDTVNRAGYVNGKQVWGTGFDRLVAADAKTGKRVWEVNTIPADASLGKQSTYVDVAVR 202
QY 137 VANGVIVASTGOYSPFGC--FVSGHDSATGEBELMRNFIPRAGEED-----182
DB 203 VANGVIVASTGOYSPFGC--FVSGHDSATGEBELMRNFIPRAGEED-----182
QY 183 -ETWGNDEARMT-----GANGQITYPEVNLVHGSTAVGPASETORGPGTLY---233
DB 261 YKTWGE--KGAVYRGGGGSTVWDSLVYDVSLLY---LAWG-----NGSPMYKRYSE 309
QY 234 --GTN-----TRPVRPDTEIYWRHQTLPROWWDOCECFEMAVTVNDVQSTEMEGQSI 287
DB 310 GISSNIFLSIALKPEGEYVWHFQATPMDQDYSVOQITLDMV-----357
QY 288 NPNAATGE--RVVLTVGPCCTGTMMQFDATGELWADRTYONMIESIDE--NGIYTVNED 345
DB 358 -----NGEMRHVIMHAP--KNGFEYVLDAKTGEBLAGKNVYQNMANGIDPLTGRPIYND 411
QY 346 AILKELDEYDVCPTFLGGRDMPAALNDSCITFIPLNNVY-----DMMMA 392
DB 412 GLYTLGKFWYIGPBLGANEHMGAYSPKTHLYLPAHQIPFGYKNOVGKFPHPDAMN 471
QY 393 VDDEFTSMYVNTSNYTKLPKCKMIGRIDAIDISTGRTLSVERAANYSPVLSTGGV 452
DB 472 VGLDMYKNGLPDPE--ARRAYIKDLHGWLLANDPYKMETVWKIDKSPMNGSVLATGDL 530
QY 453 LFNGGDIYFRALSOETGELTMOATLAVASGOAISYEVDGQYVA-----IAGG 502
DB 531 LFGGLANGEFHAYDANGSDLYKFDQSGIIASPMYTVNGQYVAVEVWGVIYPISMG 590
QY 503 GVSYGSG 509
DB 591 GVGRTSG 597

```

RESULT 4

Q92YW8 PRELIMINARY; PRT: 601 AA.

AC Q92YW8; 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.

GN KMAP.

OS Rhizobium melioli (Sinorhizobium melioli).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RM1021;

RA Fenner B.J., Tiwari R.P., Dilworth M.J.;

RT "Regulation of Cl assimilation in Sinorhizobium melioli."

RL Submitted (SEP-2000) to the EMBL/GenBank/DBD databases.

DR EMBL; AF309488; AAC31643.1;

DR HSSP; P38539; 4AAH.

DR InterPro: IPR000372; Bac_PQQ_repeat.

DR Pfam; PF01011; Bacterial_PQQ_7

DR SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 17.8%; Score 488; DB 2; Length 601;

Best Local Similarity 26.2%; Pred. No. 3.2e-27;

Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

```

QY 14 LALLAAPFAOVTVTDEL--LANPPAGEMISYQONENRHSPLQITENYGOLOLV 71
DB 8 LAIMSIGGAOVAFANDELQKLDLP--NOMALIGTDVANIARYSKLDQINDNNGKLOLVAM 66
QY 72 --ARMOGKQVQVPLIDGVYLANP--GDVIOAID--AKTGDLWEHR--ROLNINITLNS 126

```

```

DB 67 TSTGVLRGH--EGSPLVIGDLMYVHTPPNVTYALDLSKQIOYWKYEPQDPNVIPVC 125
QY 127 FGEPTRGMAI-----VANGVIVAGST-----147
DB 126 CDTVNRAGYVADNKFPLHQADITVVALDAKTGYVMSKNDATKGETNATVAPVKDI 185
QY 148 ---CQYSPFGC--FVSGHDSATGEBELMRNF-----PRAGEEDDET 184
DB 186 LVGISGGEFVGAVHATVASMADGKVLMRGYSKMGPSDPLIDPEKTHLGKFPVGDGSLTT 245
QY 185 WGNDEYARMTG--AMGQITYPEVNLVHGSTAVGPASETORGPGTLYGTFEPAV 241
DB 246 WEGD---QMKIGGTTWGYSYDPEENLVYTGTPSTWNPQR--PDNNW--SMITFAR 299
QY 242 RPDTEIYWRHQTLPROWWDOCECFEMAVTVNDVQSTEMEGQSIINPNAATGERRVLG 301
DB 300 DVTGTMAKLLXQMTFPEDEMDYGVNEMILTGHDGK-----DRKLLTH 343
QY 302 VPCKTGMQFDATGELF-----WARDTYKQNMIE-----SIDENGIV 340
DB 344 FD--RNGFGYTMQVRVTELLVAEKYDPTVMATEVMDPKSDYGRPOVVAQSTEBNG-- 400
QY 341 TVNEDAILKELDEYDVCPTFLGGRDMPAALNDSCITFIPLNNVYCYDMAVDOETSM 400
DB 401 ---EDNNTT-----GYCPALGTRKQOPAAISPKTELFYVPTNIVCMDEPFRVSYTAG 451
QY 401 DYNNTSNYTKLPKCKM--IGRIDAIDISTGRTLSVERAANYSPVLSTGGV 457
DB 452 QPYVGAATLSWYPP--KDSHGGMGNFIACDNKEKIKMSLPEPFSVSGALATAGDVFFGT 510
QY 458 TDYFRALSOETGELTMOATLAVASGOAISYEVDGQYVA-----GVSYSGLNS 512
DB 511 LEGYLAVDATGSKELYREKTPSGVIGNVYTAAREGQYVAVLSGVGAGIAGLGLTN 570

```

RESULT 5

Q92YW9 PRELIMINARY; PRT: 601 AA.

AC Q92YW9; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (BC 1.1.99.8).

GN SMH20173.

OS Rhizobium melioli (Sinorhizobium melioli).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RA MEDLINE=21396508; PubMed=11481431;

RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,

RA Vorholzer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,

RA Golding B., Puelher A.;

RT "The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-

RT fixing endosymbiont Sinorhizobium melioli."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).

DR EMBL; AL603642; CAC48573.1;

DR Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.

DR SEQUENCE 601 AA; 65759 MW; D73424EFED15ADB6 CRC64;

Query Match 17.8%; Score 487; DB 16; Length 601;

Best Local Similarity 26.2%; Pred. No. 3.8e-27;

Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

```

QY 14 LALLAAPFAOVTVTDEL--LANPPAGEMISYQONENRHSPLQITENYGOLOLV 71
DB 8 LAIMSIGGAOVAFANDELQKLDLP--NOMALIGTDVANIARYSKLDQINDNNGKLOLVAM 66

```

```

OY 72 --ARGMGRKQVQVPLIHGVWYLANP-GDVIAQID-AKTGDLIMEHR-ROLPIATLNS 126
DB 67 TESTGVLRGH-EGSPFLVIGDLMYHTPEPNTVVALDLSTKQOQVWKKEPQODERNVIVMC 125
OY 127 EEPERGMAL-----VANGVYAGST-----147
DB 126 CDTNRGVAAYDNKIEHQADITVVALDAKTGVYMSVKNKGDAIKGTNTATVMPYDKI 185
OY 148 ---COYSPFGC--FVSGHDSATGEELMRYNFI-----PRAGEGDEI 184
DB 186 LVGISGGEFVGRGHVATASMDGKVLKRGSMGSDSTLLDPEKTHLKGKPVSGKDLTT 245
OY 185 WGNDEYANWMTG---AMGQITPDVTLNVHSGTAVGASPTORSGTPTGTYGTNTREAV 241
DB 246 WEGD---OMKIGGTWGMWYSYDEENLVYGGTGNPSTWMPQR--PGDNRM-SMTIFAR 299
OY 242 RPDGEIYWRHQTLPBNDWQDECFEFEMATNVNVPSTMEGLQSLNPNATGERVLTG 301
DB 300 DVDIGMAKWLQMPHDEMDIDGVNEMILTEQOIDGK-----DRKLTH 343
OY 302 VPCRTGTMQPDATGEEL-----WARDNYONMIE-----SIDNGIV 340
DB 344 FD-RNGEGYTMDRYTGELVAEKYDPTVNWATEVYMDPKSDKGRPOVVAQYSTEONG-- 400
OY 341 TVNEDALIKELDYEDVCPTFLGGRMPSALNPDGCIPTIPLNNTGYMMAVDOETSM 400
DB 401 ---EDWTNT-----GVCPALGTRKQDQPPAAYSPKTELEVPPNHCMDYEEPRVSYTAG 451
OY 401 DVYNTSNVTKLPGRKDM---IGRIDADISTGRTLMSVERAANYSPVLSTGGGVLFNG 457
DB 452 QPYGATLTMYP-KDSHGMGNFIAMDNEKIKMSLPEPSPVMSGALATAGDVFYGT 510
OY 458 TDRFRALSOETGTLMTQTLATVASGQALSYEDVDMQYVAILAG-----GVSYSGSLNS 512
DB 511 LEGILKAVDADATGELRYKFTPSGVIGNWTVAREGQYVAVLVGVGMAIGLALJLN 570

RESULT 6
OY 046444 PRELIMINARY; PRT; 708 AA.
AC 046444;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
DE (Pc 1.1.99.?) (QH-EDH).
OS QHEDH.
ON Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 15667;
RX MEDLINE=96184549; PubMed=8654419;
RA Stoetvoegel J., Kraayveld D.E., van Stuils C.A., Jongejan J.A.,
RA De Vries S., Duine J.A.;
RT "Characterization of the gene encoding quinoahemoprotein ethanol
RT dehydrogenase of Comamonas testosteroni";
RL Eur. J. Biochem. 235:690-698(1996).
RN [2]
RP SEQUENCE OF 32-54 AND 477-490.
RX STRAIN-ATCC 15667;
RX MEDLINE=95324580; PubMed=7601151;
RA De Jong G.A.H., Geerlof A., Stoetvoegel J., Jongejan J.A., De Vries S.,
RA Duine J.A.;
RT "Quinoahemoprotein ethanol dehydrogenase from Comamonas testosteroni.
RT Purification, characterization, and reconstitution of the apoenzyme
RT with pyroloquinoline quinone analogues.";
RL Eur. J. Biochem. 230:899-905(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=86242113; PubMed=3521592;
RA Groen B.W., van Kleef M.A., Duine J.A.;

```

```

RT "Quinoahemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
RT testosteroni";
RL Biochem. J. 234:611-615(1986).
RN [4]
RP CRYSTALLIZATION.
RX MEDLINE=21536088; PubMed=11679760;
RA Oubrie A., Huizinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
RA Duine J.A., Dijkstra B.W.;
RT "Crystallization of quinoahemoprotein alcohol dehydrogenase from
RT Comamonas testosteroni: crystals with unique optical properties.";
RL Acta Crystallogr. D 57:1732-1734(2001).
CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC -1- COFACTOR: PQQ, HEME, AND CALCIUM.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
DR EMBL: X81880; CA57464.1;
DR HSP: Q924J7; 1FIC.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR003086; Cyt_C1.
DR InterPro: IPR002329; Cyt_C1C.
DR Pfam: PF01011; Bacterial_PQQ_6.
DR Pfam: PF00034; Cytochrome_c_1.
DR PRINTS: PR00605; CYTOCHROME_C1C.
KW Signal; PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
FT SIGNAL 1 31
FT CHAIN 32 708
FT TYPE 1.
FT BINDING 635 635 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 708 AA; 76822 MW; 99AB54BD06ACBACC84 CRC64;

```

Query Match 17.7%; Score 485.5; DB 2; Length 708;
 Best Local Similarity 25.7%; Pred. No. 6, 2e-27;
 Matches 152; Conservative 86; Mismatches 241; Indels 113; Gaps 20;

```

OY 3 PTSLNWSGALALLAAPPFAQVTPYT-----DELANPP-AGEMISYGQDEN 50
DB 12 PGRWMLAACLG--SAAFAQOTGPAQAQAAVQVHDSDFIRANARFPDPTGIVYAE 69
OY 51 YRSPLOTITTEVNGQLOLVARGMOPCK-VQVPLIHGVWYLANPQDVQALDATGCD 109
DB 70 TRSRIDQIDMANAVKDLGLAMSYNESTRGVEATPVVDGIMTVSASVVAHAIDTRGN 129
OY 110 LIWEHRRLP-----NIAT-----LNSFG 128
DB 130 RIWITYPQIDRSTGFKGCDDVNRGVALMKGVYGVGAMDRLLALDADATGEVHQNTF- 168
OY 129 EPTRGMAIYANG-----VIYAGSTCOYSPFGCVSGHDSATGEELMRYNFI----- 175
DB 189 EGKSGSLTTGAPRVFKGVIIIGRAEYGVAG-YITAYDDETEERKWRMSFVGDPSSK 247
OY 176 ---RAGEGDETWGNDYEARMWTCG-----WGQITPDVTLNVHSGTAVGASPTORSGT 227
DB 248 FDESMKRAARW--DPGKMWEEAGGGGTMDMSMTFDELNTMTVGVGNGSPMSHKYRSP 305
OY 228 PGGLYGTNTRAVRPDTGELVWRHQTLPBNDWQDECFEFEMATNVNVPSTMEGLQSL 287
DB 306 KGDMLYLASIVALDPDYGKYKHQETPGDNMDYTSOPMILADIKI----- 353
OY 288 NPNNAATGERRVLGVPCCTGTMQPDATGEELMARTNTONMTESTIDENG-IYTVNEDA 346
DB 354 ---AGKPKVYLHAP-KNGEFFVLDRTNKFEISAKNFVPVMAAGYDKHGKPDGI---A 405
OY 347 ILKELDYEDVCPTFLGGRMPSALNPDGCIPTIPLNNTGYMMAVDO-EFT----- 398
DB 406 AARDOSKQDANVPGRYGAHNHPRSFNPDGIVLVLPRAONVPVNLMDKKWEFQDAGPKAP 465

```

QY 399 -SMDVYNTSNVTKL--PPCKDMIGRIDAIDISTGTLSVEPAANYSPEVLSTGGGVLFNG 456
 Db 466 GSGTGWNTAKFFNABPPKSKPGGRLLAMPYAOAKAAMSVEHVSWMNGSLTTTANVVEOG 525
 QY 457 GIDRYRALSQETGELTMTOTRATATASGQALSIEYDGMQYVAIA--GGGVSYG 507
 Db 526 TADGRLLVAHAATGEKLEAPGTGYVAAPSTYMWGDKQIYSVAVNGVGYG 577

RESULT 7

Q9F9U2

PRELIMINARY; PRT: 695 AA.

ID Q9F9U2
 AC Q9F9U2
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2, Herrick J.B., Okinaka R.T., Brainerd J.B.,
 RA Terwilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2."
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: A176640; MAG09249.1;
 DR HSSP: Q9F4J7; 1F1G.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003045; CytC_heme_Dbind.
 DR InterPro: IPR003088; CytC-CI.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; Cytochrome-C; 1.
 DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN.1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 17.6%; Score 482.5; DB 2; Length 695;
 Best Local Similarity 26.8%; Pred. No. 1e-26;
 Matches 150; Conservative 81; Mismatches 228; Indels 101; Gaps 20;

QY 15 ALAAPAFQVTPYTDLL--ANPPAGEWISYGOENYRHSPLTQITTEYNGQLQLVMA 72
 Db 16 ALIYA-AGAQAQAKYDEAIRASEODGSEWLSHGRTTAEOBFSLKQIDAGNNGKGLAM 74
 QY 73 RGMOPGK-VQVTPLIHGVVYLANPQDVIAQIDAKTGDLIWEHRQLP----- 119
 Db 75 LDLENNKLEIETPLVSDGVILASISMSRVMAVADRSGKRLMOPDPQDRCHSRVTCDDAY 134
 QY 120 -NIAFLNS-----FGEPTNGMA-----IYANGYIVAG 145
 Db 135 NKGVALMNGKYYVAGALDRLIALDAKTGRELWSEQTTPAKPYSTIGAPRVVAGKYLIGN 194
 QY 146 STCCQSPFGCVSGHDSATGEELWRNYFIPRAGE-----EGDETW--GNDYEARM 193
 Db 195 GGAETGVRG-PFSAYDATETGKAMRFYIVP--GDPAPRIHPHLEAKAKTWKDDY--W 248
 QY 194 MTG----AMGOITDPTVNLVHGSYAVGPASETORGTGGTLYGINTFRVAVRPTGELY 249
 Db 249 KGGGGGYWDSMAVDPEDLLIYIGTGNGSPWNRIRSPGGGDNLYLSIIALRPDSKLL 308
 QY 250 WRRQTLPRDNDQECTFEEMANTVNDVQPTSTEMEGLQSIINPAATGSRVLTGVPCKTGM 309
 Db 309 WRRQTLPRDNDQECTFEEMANTVNDVQPTSTEMEGLQSIINPAATGSRVLTGVPCKTGM 309
 QY 310 WQFDEATGEFLWARTNYONMIESID--ENGIVYVNEADAILKELDEVDYCCPTFLGRDW 367
 Db 352 YVLDRTATGELLSAEKFKVYTAWEKVDLATGRPEVEPGSRYEKEGVVWV---PSSFGAHNV 408

QY 368 PSALNPDGSIYFIPLNNVYCYDMAVDOETSMDEVNT---SNVTKLPKCKDMT-GRID 422
 Db 409 HSMSTNPQTGLMFIYQDILPGVYRNEGATFKKIDGLTGTGTFSDTHIRP--RAVSGALL 466
 QY 423 AIDISTGTLSVEPAANYSPEVLSTGGGVLFNGTGRFALSOETGELTMTOTRATATVA 482
 Db 467 AMDPYRORAMRVPRHSFYWNGGTLSTAGNLVFOGTADQHLAHYSDKQRLMSRAQTGI 526
 QY 483 SGQALSIEYDGMQYVAIAG 502
 Db 527 VAAPISFSLDGEQYVAIVMAG 546

RESULT 8

P71509 PRELIMINARY; PRT: 601 AA.

ID P71509
 AC P71509
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MKAF.
 OS Methylobacterium extorquens.
 OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OC NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=97312011; PubMed-9168622;
 RA Christoserdova L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RT methyloctrophy gene clusters in Methylobacterium extorquens AM1."
 RL Microbiology 143:1729-1736(1997).
 DR EMBL: U72662; AAB58890.1;
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CBB239 CRC64;

Query Match 17.2%; Score 470.5; DB 2; Length 601;
 Best Local Similarity 25.9%; Pred. No. 6.1e-26;
 Matches 155; Conservative 105; Mismatches 219; Indels 119; Gaps 26;

QY 1 MKPTSLMASGALALAAPAFQVTPYTDLLANPPAGEWISYGOENYRHSPLTQIT 60
 Db 1 MRAVHLIALGAGLAA--ASPALANESVLKG--VAAP--AEQVLOTVDYANTKSKLDQIN 54
 QY 61 TENYGOQLVW--ARGMOPGVQVTPLIHGVVYLANP--GDVIAQIDAKTG--DLIWEHR- 115
 Db 55 ASNVKNIQVANTFSTGYLRKH--ESGPLVYGNIMYVHTPPFNIIYVALLDQGAIVKRYEP 113
 QY 116 RQLPNI-----ATLNSF-----GEPTNG-- 133
 Db 114 KQDSVLPVPMCCDTVNRGIAVADAILLHOADTLIVSLDAKSGKVMWSYKNGDPSKGETN 173
 QY 134 ---MAIYANGYIVAGSTQCYSPFGCVSGHDSATGEELWRNYF----- 174
 Db 174 TATVLPYKDKYIVGSGEERG--VOCHVATYDAIDKSGKVMWNGYSIGEPDQILVPEKTTSL 232
 QY 175 --PRAGEGDETWGNDYFARMMTG---AMGOITDPTVNLVHGSYAVGPASETORGTG 229
 Db 233 GKPIAKDSLSLKTWBD---QMKTGCGGCTGWGFSYDPKLDIMYGSQ--NPSYVNNPKQRP 287
 QY 230 GTLYGNTFRFAVRPDGTGIVRKHQTLPRDNDQECTFEEMANTVNDVQPTSTEMEGLQSI 289
 Db 288 DNKW--SMTIWARNDPTGAKWVYQMTPHDEWDGINDMTLTD-----OKFDS----- 334
 QY 290 NAAETGERVLTGVPCKTGMFOAETGEFLWARTN--YONMIESID--ENGIVYVNEADAI 347
 Db 335 ---KDRPLITFD--RNGFGTILDRATGEVLYAKRFDPVYVWMAKRVLDLGRGSKTYGRPLV 389

348 LKELDEYD-----VCPTEGGRDWPSPALNPDGSIYFIPLNNVCYDMAVDEPFS 399
 Db 390 VSKSTEDNGBDVNSKGCIPALCTKQOPRAISPKTGLTYFPLNHCMDYEPFRVYTP 449
 Qy 400 MDVYTSVTKLP-PGR-DMIGRIDAIDISTGRTLSVERAANYSPIVSTGGVLFNGG 457
 Db 450 GQPYVAGTSLMYPARSGHGWGNFIAMDNLQKIKMSNPQSFAMGALATSGDVFYGT 509
 Qy 458 TDRFRALSOETGELMOTRLATVAASGAISEYVDGQVAYIANG-----GVSYSGL 510
 Db 510 LEGFLKAVDSKTGKELKFKTPSGCITGNVYTHKQKHVAIVSGVAGNAGIGLAAGL 567

RESULT 9
 Q9A048 PRELIMINARY; PRT: 629 AA.
 AC Q9A048;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE MXAF.
 GN MXAF.
 OS Methylobacterium nodulans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID-114616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORS2060;
 RX MEDLINE-2056686; PubMed-11114919;
 RA Sy A., Giraud E., Jourand P., Garcia N., Willems A., de Lajudie P.,
 RA Plin Y., Neyra M., Gillis M., Boyin-Masson C., Dreyfus B.,
 RT "Methylobacterium Methylobacterium bacteria nodulate and fix nitrogen in
 RT symbiosis with legumes."
 RL J. Bacteriol. 183:214-220(2001).
 DR EMBL: AF220764; AAG49450.1;--
 DR HSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00.
 DR Pfam: PF01011; Bacterial_P00_1; 7.
 DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 SO SEQUENCE 629 AA; 68798 MW; 3DB559D8373BB0B2 CRC64;

Query Match 17.0%; Score 466.5; DB 2; Length 629;
 Best Local Similarity 24.9%; Pred. No. 1.3e-25;
 Matches 150; Conservative 98; Mismatches 214; Indels 141; Gaps 22;

4 TSLMASAGATLALAAPFAQVPTVDLL-ANPPAGEMISYGNOSNHRSPLTQTTT 62
 Db 11 TGVSYAALALALPLAGPS-----ALANDKLVLSKSDGNWMPGKNYSDNYSKLQINAE 66
 Qy 63 NVGQQLVW--ARGMQPKVOVPTLHDGVWY-----LANPGDV----- 99
 Db 67 NVKNLVKSMQESTGLLNCH-EGAPLVVDGVTYVTSFNNMTFALGDPEKIIIMQKPKQ 125
 Qy 100 -----IAIDAKTQDLWEHRRQIPNT 121
 Db 126 NPAARSVACCDLVNRGLAYWPGDCKTSLIKTLILDDGHVVALNAQGTETWIKTENDIRY 185
 Qy 122 -ATLNSFGEPTRGMAIVANGYVAGSTCOYSPFGCVSGHDSATGEELMNTYFIPRAGEE 180
 Db 186 GSTL-----TIAPYVVKDKVLISSGAEIGVNG-YLTAVDVRTGDKMAY-----ATGPD 235
 Qy 181 GD-----ETGNDYEARMKMG---WGQIYDVTNIVHYGSTA 216
 Db 236 SDLLADENITHNAHYQOKGLGTSTWGD--AMKIGGCTNMGVAYDPGTNLIYGTGN 292
 Qy 217 VGPASETORGTGGLTYGNTFAVRPDTGEIYWRHQTLPROMWDECEFFEMVYTVWDVQ 276
 Db 293 PAPNNEIMR--PGDNKV-TWTFIRARDVDYGEAKFGYOKTPEHDEMDYAGVVMWL----- 343

Qy 277 ESTEMEGLOSINPNAATGERRVLTGPCKTGTMOEDAEETGEFLMAR---DTNYOMIRS 333
 Db 344 -STQKD-----RSGKERKLLTHPDNRNGIYTTIDRTNGDILSHAKIDDY--VAVERT 391
 Qy 334 IDENGIYVNDALKELD-VEYDVCPTEFLGDMPSALNPDGSIYFIPLNNVCYDMA 392
 Db 392 VDLKSGLEVRDPEYGTBMHLARDICPSAMGYNOCHDSYDEPKLFLYMGINHCMDPE 451
 Qy 393 VDQETSMQVYNTSNVTKLPGR-----DMIGRIDAIDISTGRTLSVERAANYSPIV 446
 Db 452 FMLPYRAGQEFFVAGTLMVYGPGRGDQNMNEGLOIYATYAITGKRWEMEFVWGGTL 511
 Qy 447 STGGVLFNGSTDRYFRALSOETGELMOTRLATVAASGAISEYVDGQVAYIANGVSY 506
 Db 512 ATAGNVVFYGTLDGFTKARHSDTELMIRAKLPSGAIYPTVTHKGTQVAYI-----Y 566
 Qy 507 GSG 509
 Db 567 GVG 569

RESULT 10
 Q9AF95 PRELIMINARY; PRT: 691 AA.
 AC Q9AF95;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID-86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Varguel A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF355798; AAK27220.2;--
 SO SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 16.9%; Score 462.5; DB 2; Length 691;
 Best Local Similarity 25.9%; Pred. No. 2.8e-25;
 Matches 145; Conservative 84; Mismatches 203; Indels 127; Gaps 20;

39 GEMISYGNQENYRHSPLTQTTENVGQLOLVARGMQPK-VQVPTLHDGVMTLANPG 97
 Db 41 GEMRTHGYDAGTTRYSLAQITTPNNAKELGLVMSYDESSRGVEATPIYVDGVMYATPW 100
 Qy 98 DYIOADAKTGDLLMEHRRQIPNTATINSFGE-PTKMAIVANGYVAGSTCOYSPFGCE 156
 Db 101 SYVALADVRSKRRLMTDPEVPREKGNACCDVYNNRVA-VHEGKVFVS----- 149
 Qy 157 VSGH-----DSATGEELW-RNYFI-----PRA--GEEDETMGNDYEARMTGAW 198
 Db 150 LDGRVLAIDARTGRKWRBNTLDDDKPYITIGAPRINCKVYIANGAGFEVGRGITAY 209
 Qy 199 -----GQI-----TYDPVTNIVHYGSTAVG----- 218
 Db 210 DPTAASRGVVPGRGDSLPFEDASMEAAKTDWDPAGOVYLSGRRRHGVELDLGYRKAGF 269
 Qy 219 -----PASETORGTGGLTYGNTFAVRPDTGEIYWRHQTLPROMWDECEFFEM 268
 Db 270 CCTSAPATPSPWSHRRKSPAGGDNLTASTVALRPTGGEVVMHYOOTPADNMWDYSTODL 329
 Qy 269 MVTWVDV--OPSTEMEGLOSINPNAATGERRVLTGPCKTGTMOEDAEETGEFLMARDTN 326
 Db 330 ILADIELGGR-----RKVILHAP-KNGFEFVIDRDTGKFIASQNFV 370
 Qy 327 YQNMIESIDENGIYVNA-EDAILKELDVEYDVCPTEFLGGRDWPSPALNPDGSIYFIPLNN 385

```

Db 371 PNNMATGIDENKRIENEGAMPGLSMRF---PAPSARTWMSMSPOGLAYPEAQN 427
OY 386 VCYDMAVADEFTSMDEVNTSN-----VTKLPECKDMIGRIDAIDISTGR 430
Db 428 IFL-VLOEDKKNMS-----YNOAPGQAMAGIGMNLMLNPPRPAOPEFGRILAMPVOOK 482
OY 431 TMSVERAANSPVLSYGGGVLFGGDRFRFALSQETGFTLQTRLATVASGOATSYE 490
Db 483 EWRREHSPWNGSTLVAGNVFEGTADARLFLDARDGKELMSAPMGIVIAAPVTE 542
OY 491 VDMQVVAIA-GGGSVYGS 508
Db 543 VDGKQVSLAVGWSGVYGN 561

RESULT 11
O24759 PRELIMINARY; PRT; 633 AA.
ID 024759:
AC 024759:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN KXAF.
OS Hypomicrobium methylovorum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypomicrobium group; Hypomicrobium.
ON NCBI_Taxid=84;
RX MEDLINE=97457202; PubMed=93111140;
RC STRAIN-GM2;
RA Tanaka Y., Yoshida T., Watanabe K., Izumi Y., Mitsuaga T.;
RT Cloning and analysis of methanol oxidation genes in the methylotroph
RT Hypomicrobium methylovorum GM2.
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL; AB004097; BAA23272.1;
DR HSSP; P38539; AAAH.
DR InterPro; IPR001479; Bac_POO.
DR InterPro; IPR002372; Bac_POO_repeat.
DR Pfam; PF01011; Bacterial_POO_7.
DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
KW Signal; Oxidoreductase.
FT CHAIN 1 34 POTENTIAL.
FT SIGNAL 35 633 POTENTIAL.
SQ SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;

Query Match 15.6%; Score 428.5; DB 2; Length 633;
Best Local Similarity 24.6%; Pred. No. 7.4e-23;
Matches 146; Conservative 98; Mismatches 235; Indels 115; Gaps 22;

```

```

OY 228 PGGLYGTNTRFAVRPDTEGLVWRHQTLPDNDMDQETFEFMAWTNDVORSTEMEGLQST 287
Db 305 PGDNKW-TMTITARDADGKMKFGYOKTPHDEMDFAVYVIMSE-----QDTKE----- 353
OY 288 NPNATGERRVLYGPPCKTGTMMQFDATGEGFLA---RPTNOMTIESIDEINGIVTNE 344
Db 354 -----GKRKRLHPDRNKGIVYTLDRNGDLISADKLDDT--VVKRKHDKLSGLPYRD 405
OY 345 DAILELQVE-YDVCPFLGGRDPSALNDPSGIYFPIPLNWCYDMAVADEFTSMDEV 403
Db 406 PEFGRMDHKGEICPSAMGYHNGCHSDYPTKQLFNGINHICMDWEPMLPYRAGQFF 465
OY 404 NTSNVTCLP-PEKDM-----IGRIDAIDISTGRILMSVERAANSPVLSYGGVLENGG 457
Db 466 VGATLMYPGPKGRQNYLIGLOIKAYNAINTYKWEHMERFSWGGTTLTAGMLVYGT 525
OY 458 TDRYFRALSQETGFTLQTRLATVASGOATSYEVDMQVVAIAAG-GVSYSGL 510
Db 526 LDGFLKARNSDYGELLMKRLPSGVIGTPMYEHRKQVYIIVNMGVGVGPGVGL 579

RESULT 12
O9AGW3 PRELIMINARY; PRT; 623 AA.
ID 09AGW3:
AC 09AGW3:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
ON NCBI_Taxid=86174;
RX SEQUENCE FROM N.A.
RC SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnal A.S., Arp D.J.;
RT Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora.
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1;
DR HSSP; 092437; 1FLG.
DR InterPro; IPR002372; Bac_POO_repeat.
DR Pfam; PF01011; Bacterial_POO_4.
KW Signal.
FT CHAIN 1 29 POTENTIAL.
FT SIGNAL 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;
SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 15.5%; Score 424; DB 2; Length 623;
Best Local Similarity 25.2%; Pred. No. 1.5e-22;
Matches 159; Conservative 87; Mismatches 218; Indels 168; Gaps 26;

```

```

0Y 220 A$E$0R0T$P$-----G$Y$N$T$E$F$A$R$P$D$T$G$Y$V$K$R$O$T$L$P$R$O$M$D$O$E$T$E$M$A$T$N- 272
D$ 299 W$M$T$K$R$A$P$D$D$R$M$D$S$L$F-$T$G$O$A$Y$V$D$A$T$G$E$L$G$E$Y$O$H$P$P$M$A$M$D$E$S$G$N$S$V$Y$F$E$Y 357
0Y 273 -----V$D$O$P$T$E$M$E$G$L$O$S$I$N-----P$M$A$T$G$E$R$R$V$L$T$G$P$C$K$T$G$T$M$O$- 311
D$ 358 K$D$P$T$G$K$M$V$A$S$A$H$A$D$R$N$G$E$F$F$V$T$D$R$M$L$A$K$G$A$G$Y$P$K$P$T$S-----L$I$G$A$M$P$F$V$D$G$I$T$A$S$G 414
0Y 312 F$D$A$T$G$E$F$L$A$R$D$T$Y$N$O$N$N$I$E$S$I$D$E$N$G$I$Y$V$N$E$D$A$T$L$K$E$L$D$E$Y$V$C$P$T$F$G$S$R$D$W$P$S$A 371
D$ 415 F$D$K$T$G$K$P$I-$E$K$N$R$P$Q$P$K$S$A$K$G$S$I$Y$S-----P$P$T$L$G$T$N$H$P$M$S 458
0Y 372 L$N$D$S$G$Y$E$F$L$P$L$N$N$C$Y$D$M$A$V$D$E$F$T$S$M$O$Y$T$N$V$T$-----K$L$P$G$-----K$D$M$T 418
D$ 459 Y$S$P$D$T$G$E$F$I$P$A$H$M$A$M$D-----Y$T$E$N$T$Y$T$K$A$G$S$A$Y$L$G$G$F$R$I$K$N$L$F$D$H$V 505
0Y 419 G$R$I$D$A$I$D$I$S$T$G$R$T$L$M$S$V$E$R$A$A$N$T$P$V$L$T$G$G$V$L$F$N$G$T$D$R$Y$F$A$L$S$Q$E$T$G$E$F$L$M$O$T$L 478
D$ 506 G$I$L$R$A$I$D$P$A$R$S$L$A$O$G$R$V$P$A-$V$A$G$L$T$T$A$G$W$V$E$T$G$S$D$Y$L$K$A$F$A$K$K$R$K$L$M$K$F$O$T 564
0Y 479 A$T$V$A$G$O$A$I$S$E$V$D$M$O$V$A$V$A$G$G$S$Y$G$S$L 510
D$ 565 G$S$G$V$E$Y$V$T$M$D$E$Q$V$Y$A$L$O$G$-----Y$G$A$V 593

```

RESULT	13		
09L935			
ID	09L935	PRELIMINARY;	PRT; 599 AA.
AC	09L935;		
DT	01-OCT-2000 (TREMBLrel, 15, Created)		
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel, 19, Last annotation update)		
DE	METHANOL DEHYDROGENASE LARGE SUBUNIT.		
GN	MXAF.		
OS	Methylovorus sp. (strain SSI / DSM 11726).		
OC	Bacteria; Proteobacteria; beta subdivision; Methylophilus group;		
CC	Methylovorus.		
OX	NCBI_TaxID=81683;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SSI;		
RA	Kim Y.M.;		
RT	"Cloning and nucleotide sequence of mxaf gene of Methylovorus sp		
RT	strain SSI DSM11726 ";		
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF184915; AAD56237.2; -.		
DR	HSSP; P38539; 4AAH.		
DR	InterPro; IPR001479; Bac_PGO.		
DR	InterPro; IPR002372; Bac_PGO_repeat.		
DR	Pfam; PF01011; Bacterial_PGO_7.		
DR	PROSITE; PS00364; BACTERIAL_PGO_2; 1.		
DR	SEQUENCE 599 AA; 6513 MW; DBFF64B5D871BC91 CRC64;		

Query Match	15.38;	Score 420;	DB 2;	Length 599;
Best Local Similarity	24.5%;	Pred. No. 2.8e-22;		
Matches 143;	Conservative	98;	Mismatches 234;	Indels 108;
				Gaps 21;

```

OY      4 TSLMLASAGLLMLLAPAAQVPTPTDELLANPRAGMISYCONOENYRHSPLTOTTEN 63
Db      6 TATGFVAGLIALSAMIPIVAAAADSLKMLGANN--NNMTQGTDDYGHYSRSLQITGTG 63
OY      64 VGOLOI--VMARGMQPQKVQVPTPLIHGVAYTLAN--PGDIVAIDAKRGDILMEHR--RQ 11
Db      64 VKMLSEGLGHFTSTGLTGH--EGARVLIGDMMYINTPEPNTFALNLAEPEKILVMQHRKQ 122
OY      118 LPNIALNLNSFGSPRTGMA-----IYANQVIVA-----GSTOY 150
Db      123 DASVKAIVACCDILVNGSLAIDGGHIFKIQQLQSHGLVLAASDAKTGKELMKENCDAVGSITIQ 183
OY      151 SPP-----GC-----FVSGHDSATGGEMLRWNYFI-----PRA 177

```

[illegible]

RESULT	14			
0934G0				
ID	0934G0	PRELIMINARY:	PRT:	695 AA.
AC	0934G0;			
DT	01-DEC-2001 (TREMBLrel, 19, Created)			
DT	01-DEC-2001 (TREMBLrel, 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel, 19, Last annotation update)			
DE	LUPANINE HYDROXYLASE PRECURSOR.			
GN	LOH.			
OS	Pseudomonas sp. DH2001.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=163360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young			
RA	Rogozinski J.;			
RT	"Cloning, sequencing and analysis of the gene for Lupanine			
RT	hydroxylase, a quinocytochrome c from a Pseudomonas sp.";			
RL	Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AJ318095; CAC67410.1; .			
KW	Signal.			
FT				
FT	SIGNAL	1	26	
FT	CHAIN	27	695	LUPANINE HYDROXYLASE.
SQ	SEQUENCE	695 AA;	74956 MW;	9096C6387E437FE0 CRC64;

Query Match	15.38;	Score 419.5;	DB 2;	Length 695;
Best Local Similarity	26.08;	Pred. No. 3.8e-22;		
Matches 140;	Conservative 74;	Mismatches 220;	Indels 105;	Gaps 19;

QY	38	AGELISTYGNQENZYHNSPLTOITTEENVQLOLVYANGMO	PGKVQY	---TPLIHDGVYTLANP	96
Db	36	SGWMSLLGGGNGSOHFEALSALAKVYKNSYK	NIKGLSMETDMEAGGLGVNPLVADGY	ITGGP	95
QY	97	GDVIOADKKTGDDLWEHRROLPIATLISFG	---	EPRGAIYANGYVASTCO	---149
Db	96	PKGIYANDLKTCKNLMITYTPPEVOYDKT	SWTGWGTHVNRGLAVDDNVI	---GSYCKLLA	154
QY	150	-----YSPFGC	---	FVS	---
Db	155	VSRKTHKLWSSQSDCPKKMQATIGAPRGVCGKVF	IGNASSGPFGGDRHDLAPAKTKGH		214
QY	168	LMRNENYIPLAGE	---	EGDETGNDEYEAARM	---
Db	215	LMREYTMP	---GDPSPKPFENDLLAASAKTWGTDV	---	WKYTKGVSPMAIATYDASDTLV

QY	208	NUHNGTAAGVAGSTAGTGGTGG-----TLGTYMREAVRBDTEIYRHHOTIDPROMWQD	262
Db	249	NLFYTGSSNPAPMYMTM--FDGKNWTMTINGRD-----DTMAKMGYOKTYPHDMDF	300
QY	263	ECTFEMAYTNDVQDSTBMEGLQSLNPNAATGERRVLGVPCKTGTGMQDPAETGEFLMA	322
Db	301	AGVQMWLTLD-----QEVN-----GKMTPLSHIDRNGILYTLNRENGNIIYA	343
QY	323	BDTN-YQUMIESIDENGIVTNEAIIKELDVE-YDVCPPFLGGRDMPMAALNDQSIYF	380
Db	344	EKVDPAVAVFVKDLKGTGYPRDEFAFRMDHKNTICPSAMGHNGQSYDPESFTLY	403
QY	361	IPLNNVCYDMAVQDEFTSMDYVTFSTVTKLP-----PGKMDIGRIDADISTGRTIMSYE	436
Db	404	AGLHNHICDWEPEMLPYRAGQFEGATILAMYPGNGPTKKEMGQIRAFDLTTKAKWTKW	463
QY	437	RAAANYSPHLSTGGGVFNGSTDRYPRALSOETETLMTORLATVASGQALISYEVMQY	496
Db	464	EKFAAMGGTLYTKGVLWYATIDQYLEKALDNKQSKELMNNKMSGGIGSMYTSFPGKQY	523
QY	497	VAIAGGVSYSQSG 509	
Db	524	I-----GSMTGVG 531	

Query Match	14.3%	Score 392.5;	DB 2;	Length 573;
Best Local Similarity	24.6%	Pred. No. 2.7e-20;		
Matches 136;	Conservative 84;	Mismatches 218;	Indels 115;	Gaps 22;
OY	38	AGEMISYGNQENYRHSPLQITTEENWGLDW--ARGMOPGVQYTPLIHGDVNYL--	93	
Db	13	AGAMPATGATGGYYSQHNSPLAQINKSNKVNKAANSFSTGVLNCH-EEAPLVIDMKYIVS	71	
OY	94	ANGADVIQALIDAKTGDLWEHR-ROLNPILATLNSFGEPTKGM-----IV--ANGVIYA	144	
Db	72	AFPNNTYIALNLDPGKIYQHQRPKQDASTAKVACCDDVYDRGLAYGAGQIYKKAQNHLLA	131	
OY	145	-----GSTQCSPE-----GC-----FVSGHDSATGEEL	168	
Db	132	LDAKTGKINWEVEVCDPKYGSFLTDAPFAVDYTLVMGCSAGELGVGNVANAFLPTKTELK	191	
OY	169	WR-----NYIFRAGE--BEDETWGDNYEARMMTGA--WGLITDPVT	207	
Db	192	WRKFAFGSDSVRLAKDFRFSANPHFGQGLGTIKWEED--AAKIGGNTNWGVATDPGL	248	

Fri May 24 11:27:34 2002

walick-934-135.pep.rspt

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:01:01 ; Search time 117.28 Seconds
(without alignments)
89.973 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95

Perfect score: 498
Sequence: 1 MKPTSLWASGALLLAAP.....QPKVQVTPILHDGYMTLAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq_032802:*

- 1: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	100.0	579	19	AAW37873
2	428	85.9	579	19	AAW37874
3	404	81.1	579	19	AAW37876
4	381.5	76.6	578	19	AAW37875
5	142	28.5	754	21	AAW35987
6	131.5	26.4	738	12	AAW31993
7	131.5	26.4	738	13	AAW320192
8	127.5	25.6	443	22	ABG24430
9	126.5	25.4	742	11	AAW05235
10	122.5	24.6	948	22	ABG25909
11	122.5	24.6	1510	22	ABG21573

12	120	24.1	740	20	AAW95019
13	73.5	14.8	1229	22	ABW71736
14	72.5	14.6	343	19	AAW35001
15	71	14.3	526	22	AAW98995
16	69	13.9	711	22	ABW69743
17	66	13.3	227	22	AAW56221
18	66	13.3	1419	22	ABW52008
19	65.5	13.2	289	21	AAW42562
20	65.5	13.2	1004	22	AAW90456
21	65	13.1	43	22	AAW59378
22	65	13.1	43	22	AAW71934
23	65	13.1	824	18	AAW09614
24	64.5	13.0	255	16	AAW92829
25	64.5	13.0	1171	22	ABW64387
26	64	12.9	229	22	AAW50083
27	64	12.9	238	22	AAW90966
28	64	12.9	371	22	AAW91604
29	64	12.9	487	22	AAW50084
30	64	12.9	1016	22	AAW34349
31	63.5	12.8	93	22	AAE01744
32	63.5	12.8	115	22	AAE01694
33	63	12.7	539	22	ABW64896
34	62.5	12.6	138	22	ABG14809
35	62.5	12.6	420	16	AAW91068
36	62.5	12.6	420	18	AAW07593
37	62.5	12.6	420	21	AAW90454
38	62.5	12.6	583	20	AAW55919
39	62.5	12.6	712	22	ABW70600
40	62	12.4	577	22	ABG13760
41	61.5	12.3	73	22	ABW17647
42	61.5	12.3	221	21	AAW48548
43	61.5	12.3	1669	22	ABW64003
44	61	12.2	553	15	AAW56966
45	61	12.2	553	15	AAW56967

ALIGNMENTS

RESULT 1
AAW37873
ID AAW37873 standard; Protein: 579 AA.

AC AAW37873;

DT 10-AUG-1998 (first entry)

DE Alcohol and/or aldehyde dehydrogenase A amino acid sequence.

KW Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

KW 2-Keto-L-gulonic acid; L-ascorbic; inhibition.

OS Gluconobacter oxydans.

FT Key Location/Qualifiers
FT Peptide 1..23
FT Protein /note= "signal peptide"
FT /note= "mature protein"

PN EP832974-A2.

PD 01-APR-1998.

PF 11-SEP-1997; 97EP-0115801.

PR 19-SEP-1996; 96EP-0115001.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

Sorbitol dehydrog
Drosophila melano
Bacillus thermo
Xanthomonas sp. ch
Drosophila melano
Propionibacterium
Drosophila melano
Human ORF2326
C glutamicum prote
Human brain expres
Human bone marrow
Pathogenic Staphyl
C179FV-PP variable
Drosophila melano
Rat IgG1 Fc domain
C glutamicum prote
C glutamicum prote
TMFR-Fc fusion pro
Staphylococcus aur
Human gene 1 encod
Drosophila melano
Novel human diagno
bEGF-linker-sapori
Basic fibroblast g
bEGF/cathepsin B s
Pseudomonas fluores
Drosophila melano
Novel human diagno
Human nervous syst
Cochliobolus cardio
Drosophila melano
CC49 VL-VH-L-VH-
CC49 VL-L-VH-L-VH-

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 XX
 XX Claim 1; Pages 35-37; 59pp; English.
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 579 AA;
 SQ
 Query Match 100.0%; Score 498; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8e-49;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MKPTSLMASAGALALIAFAAQVTPYTDLLANPAGEWISYSGONENYRHSPLTQIT 60
 DB 1 mkptslmasagallalaapafagvprvdellampagewisygngenyhspitqit 60
 OY 61 TENVGQLQVWARGMQPKGVQVTPPLIHGVMYLAN 95
 DB 61 tenvgqlqvwargmqpkgvqvtplihdgvmylan 95
 RESULT 2
 AAW37874
 ID AAW37874 standard; Protein; 579 AA.
 AC AAW37874;
 XX
 XX 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 XX Gluconobacter oxydans.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 XX EP832974-A2.
 PN
 XX 01-APR-1998.
 PD
 XX 11-SEP-1997; 97EP-0115801.
 PE
 XX 19-SEP-1996; 96EP-0115001.
 PR
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA
 XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 PI WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde

PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 XX
 XX Claim 1; Pages 38-40; 59pp; English.
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 579 AA;
 SQ
 Query Match 85.9%; Score 428; DB 19; Length 579;
 Best Local Similarity 88.4%; Pred. No. 9.5e-41;
 Matches 84; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 OY 1 MKPTSLMASAGALALIAFAAQVTPYTDLLANPAGEWISYSGONENYRHSPLTQIT 60
 DB 1 mkptslmasagallalaapafagvprvdellampagewisygngenyhspitqit 60
 OY 61 TENVGQLQVWARGMQPKGVQVTPPLIHGVMYLAN 95
 DB 61 tenvgqlqvwargmqpkgvqvtplihdgvmylan 95
 RESULT 3
 AAW37876
 ID AAW37876 standard; Protein; 579 AA.
 AC AAW37876;
 XX
 XX 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 XX Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 XX Gluconobacter oxydans.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 XX EP832974-A2.
 PN
 XX 01-APR-1998.
 PD
 XX 11-SEP-1997; 97EP-0115801.
 PE
 XX 19-SEP-1996; 96EP-0115001.
 PR
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA
 XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 PI WPI: 1998-195228/18.
 DR N-PSDB: AAV29054.
 XX
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 XX Claim 1; Pages 44-46; 59pp; English.

100

CC In sequences AAC83156 and AAC83157.
XX
SQ Sequence 754 AA;

Query Match 28.5%; Score 142; DB 21; Length 754;
Best Local Similarity 40.2%; Pred. No. 1.3e-07;
Matches 35; Conservative 10; Mismatches 34; Indels 8; Gaps 2;

OY 14 LALAAPAFVAVPTVDE-----TDELLANPAGEMISYCONQENYRHSPLTQTITTT 66
DB 17 Lgcasaalafcatpvalaetgtatitnadqhpqdmmsygrtysqgrysplditkdnasn 76
OY 67 LQVWANGMQPGKVQV-TPLIHGVMY 92
DB 77 Lklawhylditnrgqegtpivdgymy 103

RESULT 6

AAR13993
ID AAR13993 standard; Protein; 738 AA.

XX AAR13993;

AC AAR13993;

DE 09-DEC-1991 (first entry)

XX A.altoacetigenes membrane-bound ADH 72KD sub-unit.

XX alcohol dehydrogenase complex; carboxylic acid production.

XX Acetobacter altoacetigenes.

XX EP448969-A.

XX 02-OCT-1991.

XX 26-FEB-1991; 91EP-0102793.

XX 26-MAR-1990; 90JP-0073440.

XX 26-FEB-1990; 90JP-0042391.

XX (NAKA-) NAKANO VINEGAR KK.

XX WPI: 1991-289462/40.

XX N-PSDB: AAQ13580.

XX Gene for membrane-bound alcohol dehydrogenase complex - obt.

XX from Acetobacter altoacetigenes, used for prodn. of enzyme for

XX converting alcohol to acid

XX Disclosure; Fig 3; 36pp; English.

CC Total DNA was prepared from A.altoacetigenes MH-24, digested with
CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
CC mixture was used to transform E.coli JM109. Probes were designed
CC based on the N-terminal amino acid sequence of the ADH complex
CC isolated from A. altoacetigenes (see AAQ13582-Q13584). The gene
CC encoding the 72KD ADH subunit was isolated and sequenced. The
CC directly sequenced N-terminal region of the purified 72KD sub-unit
CC corresponds to the sequence beginning at residue 36 of the deduced
CC sequence. This suggests that the first 35 N-terminal amino acids
CC form a leader peptide involved in secretion of the mature 72KD
CC protein. The deduced amino acid sequence has 77 per cent homology
CC with the same enzyme from A. aceti K6033. See also AAQ13581.

SQ Sequence 738 AA;

Query Match 26.4%; Score 131.5; DB 12; Length 738;
Best Local Similarity 32.6%; Pred. No. 2e-06;

Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALAAPAFVAVPTVDE-----LLANPAGEMISYCONQENYRHSPLTQTIT 61
DB 17 tagticaaalsysatmasadqggaatgeaallhaddhpnwmtygrtysdqyspldqtnr 76
OY 62 ENVGQLQVWANGMQPGKVQV-TPLIHGVMY 92
DB 77 snvgnlklawylditnrgqegtpivdgymy 108

RESULT 7

AAR20192
ID AAR20192 standard; Protein; 738 AA.

XX AAR20192;

DE 16-APR-1992 (first entry)

XX ADH complex protein (mol.wt. 72.000).

XX Alcohol dehydrogenase; acetic acid; fermentation.

XX Acetobacter altoacetigenes NH-24.

XX JF0326988-A.

XX 27-NOV-1991.

XX 26-MAR-1990; 90JP-0073440.

XX 26-FEB-1990; 90JP-0042301.

XX 26-MAR-1990; 90JP-0073440.

XX (NAKA-) NAKANO SUTEN KK.

XX WPI: 1992-019325/03.

XX N-PSDB: AAQ20383.

XX Alcohol dehydrogenase complex structural gene - used in plasmid

XX PT and enhancing efficiency of acetic acid fermentation for

XX PT transformed acetic acid bacteria

XX Disclosure; Fig 3(1-3); 21pp; Japanese.

XX Acetobacter transformed with the sequence encoding this protein can

XX enhance the efficiency of acetic acid fermentation. The ADH complex

XX can be easily extracted from the bacteria and purified and it can be

XX used for the determination of an alcohol.

XX See also AAQ20383-84, and -86-88.

XX SQ Sequence 738 AA;

Query Match 26.4%; Score 131.5; DB 13; Length 738;
Best Local Similarity 32.6%; Pred. No. 2e-06;
Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

RESULT 8

ABG24430
ID ABG24430 standard; Protein; 443 AA.

XX ABG24430;

XX

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #24421.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS88617.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 54789; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridization probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptides and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 443 AA:

Query Match 25.6%; Score 127.5; DB 22; Length 443;
Best Local Similarity 33.7%; Pred. No. 3.1e-06;
Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LLMAS-----AGALLAAPAFQVTPYTDLLANPAGWISYGOENVRHSPLTQ 58
DB 168 ltwgfdngpgeingtltsadapaea-lspvadt-----dwpaygrngqgrfslpq 218
OY 59 ITTENWGQLOLVMA-----RGMOQKV--QVPLIHDCVMYL 93
DB 219 inadhvnlkkaewfrtgdkvqndpgeitnevrplkvgtllyl 262

RESULT 9
AAR05235
ID AAR05235 standard; protein: 742 AA.

XX AAR05235;

XX 04-AUG-1990 (first entry)

DE Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
DE 1-2229 of alcohol dehydrogenase (ADH) gene.

XX Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;

XX Acetobacter aceti K1006 (FERM-7528); plasmid PAA721.

XX Acetobacter aceti strain K1006 (FERM-7528).

XX JP02000452-A.

XX 05-JAN-1990.
XX 30-OCT-1987; 87JP-0273190.

XX 30-OCT-1987; 87JP-0273190, JP-075069.

XX (QPPP) QP CORP (QPPJ-).

XX WPI; 1990-047990/07.

XX N-PSDB; AAQ91811.

XX Cloning DNA, plasmid and microbe contg. it
PT contains alcohol dehydrogenase gene derived from Acetobacter
PT aceti K1006 strain
XX
XX Disclosure; Fig 4; 8pp; Japanese.

XX Also new are a recombinant plasmid contg. its encoding DNA, and a
XX microorganism transformed with the plasmid.
XX A DNA fragment was obtd. from A. aceti
XX strain K1006 and was transferred to an ADH-defective strain. The
XX resulting plasmid, PAA721, contg. its gene was inserted into A.
XX pasteurianus strain IFO 3191 by the triparent method. Its gene is
XX useful for improving Acetobacter culturing.

SQ Sequence 742 AA:

Query Match 25.4%; Score 126.5; DB 11; Length 742;
Best Local Similarity 38.8%; Pred. No. 7.7e-06;
Matches 33; Conservative 11; Mismatches 38; Indels 3; Gaps 2;

OY 11 AGALLAAPAFQVTPYTDLL--ANPAGWISYGOENVRHSPLTQITTENWGQLO 68
DB 23 aaalpyaavparaagqgnlgeallhaddhpenwisrygrysqtryspldqdnrsvglk 82
OY 69 LVWARGMOQKVQ-VNPLIHDCVMY 92
DB 83 llygyltdnrggeatpplvvdgimy 107

RESULT 10

ABG25909
ID ABG25909 standard; protein: 948 AA.

XX ABG25909;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #25900.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS90096.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 XX biodiversity
 XX
 PS Claim 20; SEQ ID No 56268; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 948 AA;
 Query Match 24.6%; Score 122.5; DB 22; Length 948;
 Best Local Similarity 34.4%; Pred. No. 3.1e-05;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;
 QY 6 LWMAS-----AGATALLAARFAQVTPVTDLLANPPAGEMISYGONENYRHSPLTQ 58
 Db 310 ltwagfndpgeingtlisadatpaea-ispvadq-----dwpaygrngsgqgfspklq 360
 QY 59 ITTENVGOLQLVWA-----RGMPGKV--QVTP 85
 Db 361 inadvnhlkeawvftgtvdkpndpgeltnevtpl 396
 RESULT 11
 ABG21573
 ID ABG21573 standard; Protein; 1510 AA.
 XX
 XX ABG21573;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21564.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.

XX
 PN W0200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS85760.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 XX biodiversity
 XX
 PS Claim 20; SEQ ID No 51932; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 1510 AA;
 Query Match 24.6%; Score 122.5; DB 22; Length 1510;
 Best Local Similarity 34.4%; Pred. No. 5.6e-05;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;
 QY 6 LWMAS-----AGATALLAARFAQVTPVTDLLANPPAGEMISYGONENYRHSPLTQ 58
 Db 310 ltwagfndpgeingtlisadatpaea-ispvadq-----dwpaygrngsgqgfspklq 360
 QY 59 ITTENVGOLQLVWA-----RGMPGKV--QVTP 85
 Db 361 inadvnhlkeawvftgtvdkpndpgeltnevtpl 396
 RESULT 12
 AAW95019
 ID AAW95019 standard; Protein; 740 AA.
 XX
 XX AAW95019;
 XX
 DT 21-MAY-1999 (first entry)
 XX
 DE Sorbitol dehydrogenase (SUDH) protein sequence.
 XX
 KW Sorbitol dehydrogenase; SUDH; open reading frame; ORF2; L-sorbitose;

•
•
• Fri May 24 11:27:21 2002

us-08-934-506a-5_copy_1_95.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:59:06 ; Search time 44.35 Seconds
(without alignments)
52.321 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPTSLMASGALALIAAP.....QPGKVVTPLIHGWTILAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	28.5	754	4 US-09-296-284-4	Sequence 4, Appl
2	133	26.7	720	4 US-09-286-284-25	Sequence 25, Appl
3	131.5	26.4	738	1 US-07-985-458-3	Sequence 3, Appl
4	120	24.1	739	3 US-07-136-251-2	Sequence 2, Appl
5	61	12.2	275	3 US-08-463-903-8	Sequence 8, Appl
6	61	12.2	275	4 US-07-935-695-8	Sequence 8, Appl
7	61	12.2	280	3 US-08-463-903-10	Sequence 10, Appl
8	61	12.2	280	4 US-07-935-695-10	Sequence 10, Appl
9	61	12.2	282	3 US-08-463-903-12	Sequence 12, Appl
10	61	12.2	282	4 US-07-935-695-12	Sequence 12, Appl
11	61	12.2	553	2 US-08-263-911-7	Sequence 7, Appl
12	61	12.2	553	2 US-08-263-911-9	Sequence 9, Appl
13	60.5	12.1	587	2 US-08-436-664-23	Sequence 23, Appl
14	60.5	12.1	587	3 US-09-135-642-23	Sequence 23, Appl
15	60.5	12.1	587	3 US-08-394-232a-23	Sequence 23, Appl
16	60.5	12.1	587	3 US-08-394-232a-23	Sequence 23, Appl
17	60.5	12.1	588	5 US-08-436-664-25	Sequence 25, Appl
18	60.5	12.1	588	3 US-09-135-642-25	Sequence 25, Appl
19	60.5	12.1	588	3 US-08-394-232a-25	Sequence 25, Appl
20	60.5	12.1	588	5 PCT-US95-04080-25	Sequence 25, Appl
21	60.5	12.1	590	2 US-08-436-664-27	Sequence 27, Appl
22	60.5	12.1	590	3 US-09-135-642-27	Sequence 27, Appl
23	60.5	12.1	590	3 US-08-394-232a-27	Sequence 27, Appl
24	60.5	12.1	590	3 PCT-US95-04080-27	Sequence 27, Appl
25	60.5	12.1	624	3 US-09-057-969-3	Sequence 3, Appl
26	60.5	12.1	876	2 US-08-436-664-20	Sequence 20, Appl
27	60.5	12.1	876	2 US-08-436-664-32	Sequence 32, Appl

28	60.5	12.1	876	2 US-08-436-664-34	Sequence 34, Appl
29	60.5	12.1	876	3 US-09-135-642-20	Sequence 20, Appl
30	60.5	12.1	876	3 US-09-135-642-32	Sequence 32, Appl
31	60.5	12.1	876	3 US-09-135-642-34	Sequence 34, Appl
32	60.5	12.1	876	3 US-08-394-232a-20	Sequence 20, Appl
33	60.5	12.1	876	3 US-08-394-232a-32	Sequence 32, Appl
34	60.5	12.1	876	3 US-08-394-232a-34	Sequence 34, Appl
35	60.5	12.1	876	5 PCT-US95-04080-20	Sequence 20, Appl
36	60.5	12.1	876	5 PCT-US95-04080-32	Sequence 32, Appl
37	60.5	12.1	876	5 PCT-US95-04080-34	Sequence 34, Appl
38	60.5	12.1	954	3 US-09-057-969-2	Sequence 2, Appl
39	60	12.0	251	4 US-08-858-207a-354	Sequence 354, Appl
40	59.5	11.9	291	4 US-09-286-691-28	Sequence 28, Appl
41	59.5	11.9	291	4 US-09-687-147-28	Sequence 28, Appl
42	59	11.8	176	4 US-09-213-343-3	Sequence 3, Appl
43	59	11.8	176	4 US-09-413-814-85	Sequence 85, Appl
44	59	11.8	223	2 US-08-121-436a-2	Sequence 2, Appl
45	59	11.8	568	4 US-09-160-065-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ. ID NOS: 87
; SOFTWARE: Patentln Ver. 2.0
; SEQ. ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
US-09-296-284-4
```

Query Match 28.5%; Score 142; DB 4; Length 754;
Best Local Similarity 40.2%; Pred. No. 9.2e-09;
Matches 35; Conservative 10; Mismatches 34; Indels 8; Gaps 2;

```
QY 14 LALLAAPAFQVTPY-----TDELLANPPAGEMISYGNQENTRHSPLTOTTENWGO 66
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 17 LGCAMALAFCAISPVALAEDGTATNADQHPGDMSYGRYSBORSYPLDQITKNASN 76
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 67 LQLVWARGNQPKQVY-TPLIHDCVMY 92
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 77 LKLANHYDLDTNRGEGTPLLVDGMY 103
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 2
US-09-296-284-25
; Sequence 25, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ. ID NOS: 87
; SOFTWARE: Patentln Ver. 2.0
```


SOFTWARE: MS-Word for Windows, Ver. 7.0
SEQ ID NO 10
LENGTH: 280
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFV UHM5.8
LOCATION: 1..280
OTHER INFORMATION: :
US-07-935-695-10

Query Match
Best Local Similarity 12.2%; Score 61; DB 4; Length 280;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

QY 9 ASAGALALLAAPAFQVT-----PVT--DELLANPPAGEMISYGNQENY----- 51
DB 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSLISGNQKNYLAWYQOK 67
QY 52 -RHSPLTQITTEENVGQLQVWARGMQPG 78
DB 68 PGQSP-----KLLIYWASARESG 85

RESULT 9
US-08-463-903-12
Sequence 12, Application US/08463903
Patent No. 6071515
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/08/463,903
CURRENT FILING DATE: 1995-06-05
EARLIER FILING DATE: 1992-08-21
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-Word for Windows, Ver. 7.0
SEQ ID NO 12
LENGTH: 282
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFV UHM5.2
LOCATION: 1..282
US-08-463-903-12

Query Match
Best Local Similarity 12.2%; Score 61; DB 3; Length 282;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

QY 9 ASAGALALLAAPAFQVT-----PVT--DELLANPPAGEMISYGNQENY----- 51
DB 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSLISGNQKNYLAWYQOK 67
QY 52 -RHSPLTQITTEENVGQLQVWARGMQPG 78
DB 68 PGQSP-----KLLIYWASARESG 85

RESULT 10
US-07-935-695-12
Sequence 12, Application US/07935695
Patent No. 6329507
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.

APPLICANT: Kottle, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/07/935,695
CURRENT FILING DATE: 1992-08-21
PRIOR APPLICATION NUMBER: US 08/463,903
PRIOR FILING DATE: 1995-06-05
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-Word for Windows, Ver. 7.0
SEQ ID NO 12
LENGTH: 282
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFV UHM5.2
LOCATION: 1..282
OTHER INFORMATION: :
US-07-935-695-12

Query Match
Best Local Similarity 12.2%; Score 61; DB 4; Length 282;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

QY 9 ASAGALALLAAPAFQVT-----PVT--DELLANPPAGEMISYGNQENY----- 51
DB 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSLISGNQKNYLAWYQOK 67
QY 52 -RHSPLTQITTEENVGQLQVWARGMQPG 78
DB 68 PGQSP-----KLLIYWASARESG 85

RESULT 11
US-08-263-911-7
Sequence 7, Application US/08263911
Patent No. 5877291
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S
APPLICANT: Gourlie, Brian B
TITLE OF INVENTION: MULTIVALENT SINGLE CHAIN ANTIBODIES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Duane C. Ulmer
STREET: P.O. Box 1967
CITY: Midland
STATE: MI
COUNTRY: US
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/263,911
FILING DATE: 21-JUN-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,263
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ulmer, Duane C
REGISTRATION NUMBER: 34,941
REFERENCE/DOCKET NUMBER: C-41,014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 636-8104
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear


```
Patent No. 6066483
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/135,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/394,232
FILING DATE:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-09-135-642-23

Query Match          12.1%; Score 60.5; DB 3; Length 587;
Best Local Similarity 38.3%; Pred. No. 42;
Matches 18; Conservative 5; Mismatches 21; Indels 3; Gaps 1;

OY 17 LAAPAFQVTPVDELANPPAGEMISYGNQENYRHSPLTOTTEN 63
|| || |||||:|:| | || | | | | | | | | | | | | | |
Db 7 LAGMDFALADSVDLMDLADKALVVEYV---DNYHAPIVGIALAN 50

RESULT 15
US-08-394-232a-23
Sequence 23, Application US/08394232A
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/135,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/394,232
FILING DATE:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-09-135-642-23
```

```
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-394-232a-23
```

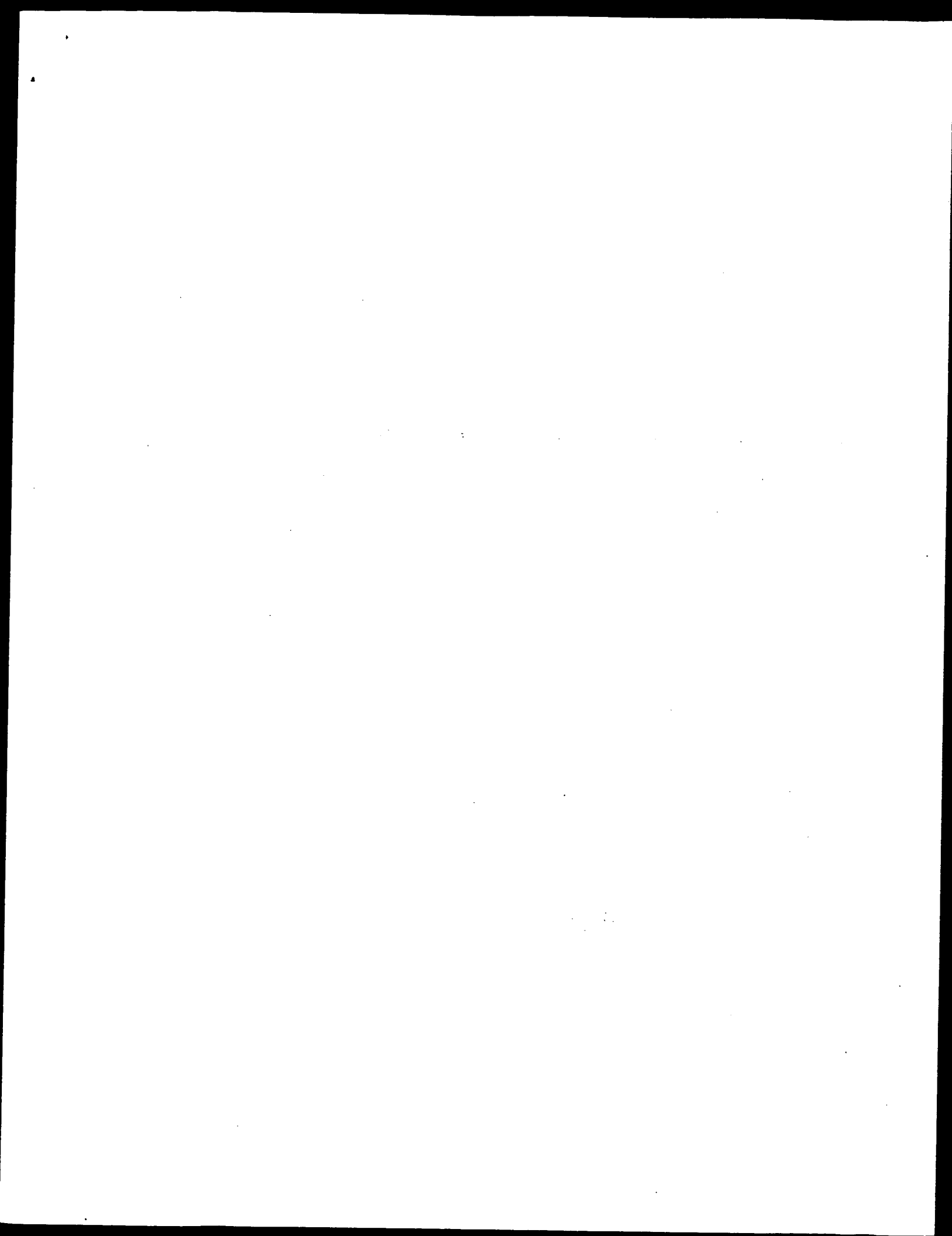
```
Query Match          12.1%; Score 60.5; DB 3; Length 587;
Best Local Similarity 38.3%; Pred. No. 42;
Matches 18; Conservative 5; Mismatches 21; Indels 3; Gaps 1;

OY 17 LAAPAFQVTPVDELANPPAGEMISYGNQENYRHSPLTOTTEN 63
|| || |||||:|:| | || | | | | | | | | | | | | | |
Db 7 LAGMDFALADSVDLMDLADKALVVEYV---DNYHAPIVGIALAN 50
```

Search completed: May 24, 2002, 10:02:54
Job time: 228 sec

Fri May 24 11:27:21 2002

us-08-934-506a-5_copy_1_95.ra1



```

Query Match      28.0%, Score 139.5, DB 2, Length 742;
Best local Similarity 40.0%, Pred. No. 5, le-07;
Matches 34; Conservative 10; Mismatches 38; Indels 3; Gaps 2;

Oy      11 AGGALAAPAAQVTPVDEL--ANPAGHEWISYGOQENRHSPLPTQTTEENVQQLQ 68
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      23 AALPLPAAAPARADGGNTGEALIHADDPENWLISYGRYSEORYSPLDQINRSNVGLK 82

Oy      69 LVMARGMQPGKYQ-VTPILHGVMT 92
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      83 LAMYYTLDTNRGGEATPLVVDGIMY 107

RESULT      2
S14270      alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
C:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S14270
R:Ramak, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A>Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:q216185; PIDN:BA00528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <Sig>

```

F;36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAT>

Query Match 26.4%; Score 131.5; DB 2; Length 738;
Best Local Similarity 32.6%; Pred. No. 3.8e-06;

Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALAAPAFQAVPTVDE-----LLANPPAGEWISYQONENYHSPPLQITTT 61
DB 17 TAGTTCALISGYATWASADSGATGEGAIHADHPENMMWTGRTYSQDRTSPLDQINR 76
OY 62 ENVGOLVWARGMOPGRKQV-VPLIHGVMY 92
DB 77 SNVGNLKLAWYLDLTNRGQESTPLVDGVMY 108

RESULT 3

AG0523

glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typh1 (st

C:Species: Salmonella enterica subsp. enterica serovar Typh1

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence.revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AG0523

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AG0523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:916501455; GSPDB:GN00176

C:Genetics:

A:Gene: STR0191

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 26.0%; Score 129.5; DB 2; Length 796;
Best Local Similarity 34.6%; Pred. No. 6.8e-06;

Matches 36; Conservative 14; Mismatches 29; Indels 25; Gaps 5;

OY 6 LWAS-----AGALALAAPAFQAVPTVDELLANPPAGEWISYQONENYHSPPLQ 58
DB 135 LTMAGFNDPQINGTLRADATPA-ATSSSIAD-----DPAVGRNGEGGRYSPLKQ 185
OY 59 ITTENVGOLVWV-----RCMOPGRK--CVTPLIHGVWYL 93
DB 186 ITADNVHQLKEAVVFTGDLKQPNDEGTITNEVTPIKGDIYL 229

RESULT 4

JS0326

alcohol dehydrogenase (EC 1.1.1.1) 72k chain precursor - Acetobacter acetii

C:Species: Acetobacter acetii

C:Date: 31-Mar-1990 #sequence.revision 31-Mar-1990 #text_change 20-Jun-2000

C:Accession: JS0326

R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.

J. Bacteriol. 171, 3115-3122, 1989

A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase sub

A:Reference number: JS0326; MUID:89255070

A:Accession: JS0326

A:Molecule type: DNA

A:Residues: 1-742 <INO>

A:Cross-references: GB:D90004; GB:M26951; NID:g216193; PIDN:BA14058.1; PID:g216194

A:Experimental source: strain K6033

A:Note: amino terminal of mature protein is confirmed

C:Genetics:

A:Gene: adh1

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

F;135/Domain: signal sequence #status predicted <SIG>

F;36-742/Product: alcohol dehydrogenase 72k dehydrogenase chain #status predicted <AD

Query Match 25.6%; Score 127.5; DB 2; Length 742;
Best Local Similarity 38.8%; Pred. No. 1e-05;

Matches 33; Conservative 11; Mismatches 38; Indels 3; Gaps 2;

OY 11 AGALALAAPAFQAVPTVDELL--ANPPAGEWISYQONENYHSPPLQITTTENYGO 68
DB 23 AAALPYAAVPAADQGGTGEALIHADHPENMLSTGRYSQDRTSPLDQINRNVGDLK 82
OY 69 LVMAGMOPGRKQV-VPLIHGVMY 92
DB 83 LIGYITLDTNRGQESTPLVDGIMY 107

RESULT 5

JV0107

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli

C:Species: Escherichia coli

C:Date: 30-Jun-1992 #sequence.revision 12-Dec-1997 #text_change 16-Jun-2000

C:Accession: D64735; JV0107; A45997; S45201; I41228

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.D.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: D64735

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-796 <BLAT>

A:Cross-references: GB:AE000122; GB:U00096; NID:91786315; PIDN:AMC73235.1; PID:g17863

A:Experimental source: strain K-12, substrain M61655

R:Cleton-Jansen, A.M.; Goosen, M.; Fayet, O.; van de Putte, P.

J. Bacteriol. 172, 6308-6315, 1990

A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino

A:Reference number: JV0107; MUID:91035240

A:Accession: JV0107

A:Molecule type: DNA

A:Residues: 1-58, 'L', 60-148, 'H', 150, 'K', 194-796 <CLE>

A:Cross-references: GB:X51323; NID:941553; PIDN:CA55706.1; PID:941554

A:Experimental source: strain K12

R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.

J. Biol. Chem. 268, 12812-12817, 1993

A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia co

A:Reference number: A45997; MUID:93286127

A:Accession: A45997

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <YAM>

R:Fujita, N.

submitted to the EMBL Data Library, January 1994

A:Reference number: S45181

A:Accession: S45201

A:Molecule type: DNA

A:Residues: 1-796 <FUJ>

A:Cross-references: EMBL:D26562; NID:9473770; PIDN:BA05580.1; PID:g473791

A:Experimental source: strain K-12 substrain W3110

R:Yamada, M.; Asaka, S.; Saiter, M.H.; Yamada, Y.

J. Bacteriol. 175, 568-571, 1993

A:Title: Characterization of the gcd gene from Escherichia coli K-12 W3110 and regula

A:Reference number: I41228; MUID:93123180

A:Accession: I41228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148, 'H', 150, 'K', 194-665, 'H', 667-796 <RES>

A:Cross-references: GB:D12651; NID:9216555; PIDN:BA02174.1; PID:g216556

C:Genetics:

A:Gene: gcd

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:11-37/Domain: transmembrane #status predicted <TM1>
 F:41-39/Domain: transmembrane #status predicted <TM2>
 F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93/95/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:466/Active site: Asp #status predicted

Query Match 25.6%; Score 127.5; DB 1; Length 796;
 Best Local Similarity 33.7%; Pred. No. 1.1e-05;
 Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LMAS-----AGALALLAPAFQVTPVDELNPAGEMISYSGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTISADATPAEA-ISPVDQ-----DMPAYGRNQGGRFSPKQ 185

OY 59 ITTENVGQLQVWA-----RGMPGRV--QVTPLIHDSVMTL 93
 DB 186 INADNVHNLKEAWFRTGDVKQPNDEITNEVTPIKVGDITLXL 229

RESULT 6
 H90644
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A96629; MUID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA33551.1; PID:G13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 050952
 C:Genetics:
 A:Gene: EC50128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 25.6%; Score 127.5; DB 2; Length 796;
 Best Local Similarity 33.7%; Pred. No. 1.1e-05;
 Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LMAS-----AGALALLAPAFQVTPVDELNPAGEMISYSGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTISADATPAEA-ISPVDQ-----DMPAYGRNQGGRFSPKQ 185

OY 59 ITTENVGQLQVWA-----RGMPGRV--QVTPLIHDSVMTL 93
 DB 186 INADNVHNLKEAWFRTGDVKQPNDEITNEVTPIKVGDITLXL 229

RESULT 7
 H85495
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: AB5480; MUID:21074935; PMID:11206551
 A:Accession: H85495
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-796 <STO>
 A:Cross-references: GB:AE005174; MID:912512839; PIDN:AG54428.1; GSPDB:GN00145; UNGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 25.6%; Score 127.5; DB 2; Length 796;
 Best Local Similarity 33.7%; Pred. No. 1.1e-05;
 Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LMAS-----AGALALLAPAFQVTPVDELNPAGEMISYSGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTISADATPAEA-ISPVDQ-----DMPAYGRNQGGRFSPKQ 185

OY 59 ITTENVGQLQVWA-----RGMPGRV--QVTPLIHDSVMTL 93
 DB 186 INADNVHNLKEAWFRTGDVKQPNDEITNEVTPIKVGDITLXL 229

RESULT 8
 S52317
 C:Species: Comamonas testosteroni
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S52317; S52317; S52317
 R:Stoerovogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Du
 Eur. J. Biochem. 235, 690-698, 1996
 A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenas
 A:Reference number: S52366; MUID:96184549
 A:Accession: S52366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STO1>
 A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CA57464.1; PID:g663196
 A:Accession: S52373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STO2>
 R:de Jong, G.A.H.; Geerlof, A.; Stoerovogel, J.; Jongejan, J.A.; de Vries, S.; Duine,
 Eur. J. Biochem. 230, 899-905, 1995
 A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific
 A:Reference number: S55908; MUID:95324580
 A:Accession: S55908
 A:Molecule type: protein
 A:Residues: 32-50, 'X', 52-54; 477-483, 'X', 485-490 <DEJ>
 A:Experimental source: ATCC 15667
 C:Genetics:
 A:Gene: qhdh
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; quinoprotein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 23.8%; Score 118.5; DB 2; Length 708;
 Best Local Similarity 30.5%; Pred. No. 9.2e-05;
 Matches 32; Conservative 18; Mismatches 40; Indels 15; Gaps 4;

OY 3 PTSLMASGALALLAPAFQVTPV-----DELLANP-AGEWISYSGONEN 50
 DB 12 PGRWMLLAACIG--SAAAFQOTGPAAQAAAVORVDDFTIRANNAKTPDPPTIGVDYAE 69
 OY 51 YRHSPLQITTENVGQLQVWARGMPGRV--QVTPLIHDSVMTL 94
 DB 70 TRYSLKDQINANVMDLGLAMSYNLESTRGVGEATPVVVDGIMYVS 114

RESULT 9
 G98221
 C:Species: Agrobacterium tumefaciens (strain C58, Cere
 glucose dehydrogenase PA2290 [Imported] - Agrobacterium tumefaciens (strain C58, Cere

Query Match 22.4%; Score 111.5; DB 2; Length 778;
Best Local Similarity 28.7%; Pred. No. 0.00059;
Matches 35; Conservative 15; Mismatches 39; Indels 33; Gaps 5

[illegible]

Query Match	21.8%	Score 108.5	DB 2	Length 803
Best Local Similarity	27.5%	Pred. No. 0.0013		
Matches	36	Conservative 14	Mismatches 34	Indels 47
			Gaps	5
QY	6	LLM-----ASAGALALALAAPAFQAVTVPTDELLANP-----		36
Db	103	LLMTPMRRLPLDGNAPLCTAALGAAVYLAGAAAVSQ-FTNGQIVYGRIDRDSGTST		161
QY	37	-----PAGEWISTYQONCENTRHSPFLQITTEYNGQQLTWAMRGMPGVY-----		81
Db	162	APAWPDODMOWAYGRTEFGDRYSPLKQIPANVGOLEEAMR-INTGDLPPAADPDELTNE		219
QY	82	VTPLIHGVMY 92		
Db	220	NTPLKNGMLY 230		

RESULT 12
OPREX
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxydans
C:Species: Gluconobacter oxydans
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_Change 22-May-1998
C:Accession: S17716; S19265
R:Clatton-Jensen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
MOL: Gen. Genet. 229, 206-212, 1991
A:Title: A single amino acid substitution changes the substrate specificity of quinop
A:Reference number: S17716; MUID:92017653
A:Accession: S17716
A:Molecule type: DNA
A:Residues: 1-808 <CLE>
A:Cross-references: EMBL:X62710
R:Goosen, N.
submitted to the EMBL Data Library, February 1992
A:Reference number: S19265
A:Accession: S19265
A:Molecule type: DNA
A:Residues: 1-212/'A',214-808 <GOO>
A:Cross-references: EMBL:X62710; NID:958416; PID:958417
C:Genetics:
A:Gene: gdh
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

QY 10 SAGALLA-AP-AFAQVTPVTDELNPAGE-WISYQONENYRHSPLTOITENYGO 66
 Db 7 SVSALAMIALAPALSSGAYANDKVELSKSDDNWVMPGKNYDSNNEFDLKQINKGNVYKQ 66
 QY 67 LQLYW--ARGMOPGKYQVTPPLIHGVMYL 93
 Db 67 LRPAMTFSTGLNGH-EGAPLYVDGKMYI 94

Search completed: May 24, 2002, 10:06:02
 Job time: 286 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:03:02 ; Search time 27.86 Seconds
(without alignments)
132.030 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPTSLMASAGALALAAP.....QPKGVQVTPILHGWATLAN 95

Scoring table: BIOSIM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	158.5	31.8	757	1	DHET_GLUOX
2	131.5	26.4	738	1	DHET_ACERO
3	130.5	26.2	739	1	DHET_ACERO
4	127.5	25.6	742	1	DHET_ACERO
5	127.5	25.6	742	1	DHET_ACERO
6	108	21.7	808	1	DHET_GLUOX
7	104	20.9	626	1	DHET_GLUOX
8	95	19.1	801	1	DHET_GLUOX
9	94	18.9	626	1	DHET_GLUOX
10	93	18.7	631	1	DHET_GLUOX
11	91	18.3	631	1	DHET_GLUOX
12	76	15.3	733	1	DHET_GLUOX
13	76	15.3	733	1	DHET_GLUOX
14	72.5	14.6	776	1	DHET_GLUOX
15	70.5	14.2	600	1	DHET_GLUOX
16	68.5	13.8	417	1	DHET_GLUOX
17	65.5	13.2	661	1	DHET_GLUOX
18	65.5	13.2	661	1	DHET_GLUOX
19	64	12.9	325	1	DHET_GLUOX
20	64	12.9	325	1	DHET_GLUOX
21	63.5	12.8	369	1	DHET_GLUOX
22	63.5	12.7	369	1	DHET_GLUOX
23	62.5	12.6	128	1	DHET_GLUOX
24	62.5	12.6	128	1	DHET_GLUOX
25	62	12.4	310	1	DHET_GLUOX
26	62	12.4	310	1	DHET_GLUOX
27	62	12.4	310	1	DHET_GLUOX
28	62	12.4	310	1	DHET_GLUOX
29	61.5	12.3	4367	1	DHET_GLUOX
30	61.5	12.3	4367	1	DHET_GLUOX
31	61	12.2	221	1	DHET_GLUOX
32	61	12.2	221	1	DHET_GLUOX
33	61	12.2	221	1	DHET_GLUOX

34	61	12.2	753	1	METE_SALTY
35	61	12.2	767	1	HYFE-SYN3
36	61	12.2	864	1	STAZ_PIG
37	60.5	12.1	876	1	DPOL_BACST
38	59.5	11.9	139	1	RS11_BACST
39	59.5	11.9	139	1	RS11_BACST
40	59.5	11.9	146	1	ADUP_MLCTA
41	59.5	11.9	221	1	SEGA_BPT4
42	59	11.8	176	1	RIMG_ASPEP
43	59	11.8	233	1	RLI_SERMA
44	59	11.8	270	1	CB12_LYCES
45	59	11.8	300	1	MOXI_METEX

ALIGNMENTS

RESULT ID	DHET_GLUOX	STANDARD	PRT	757 AA
AC	005542			
DE	01-NOV-1997 (Rel. 35, Created)			
DE	01-NOV-1997 (Rel. 35, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).			
GN	ADHA			
OS	Glucobacter oxydans (Glucobacter suboxydans).			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae; Glucobacter.			
OX	NCBI_TaxID=442;			
RM	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.			
RC	STRAIN-IFO 12528;			
RX	MEDLINE=97208225; PubMed=9055427;			
RA	Kondo K., Horinouchi S.;			
RT	"Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Glucobacter suboxydans and their expression in Acetobacter pasteurianus."			
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).			
RL	FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).			
CC	(BY SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.			
CC	-1- COFACTOR: PO4 AND HEME.			
CC	-1- SUBUNIT: HETERODIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL PO4 DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).			
CC	EMBL: D86375; BAA19753.1; -			
DR	HSSP: Q97437; 1F1G.			
DR	InterPro: IPR001479; Bac_P00.			
DR	InterPro: IPR002372; Bac_P00.repeat.			
DR	InterPro: IPR000345; CytC_heme_bind.			
DR	Pfam: PF01011; Bacterial_P00_1; 1.			
DR	PROSITE: PS00363; BACTERIAL_P00_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_P00_2; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C; 1.			
DR	Oxidoreductase; PO4; Heme; Periplasmic; Membrane; Signal.			
KW	SIGNAL			
FT	CHAIN			
FT	MOD_RSS			
FT	ACT_SITE			

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA: 82968 MW: 3989F90E3B947581 CRC64.

Query Match 31.8%; Score 158.5; DB 1; Length 757;
 Best Local Similarity 45.7%; Pred. No. 3.5e-09;
 Matches 42; Conservative 12; Mismatches 31; Indels 7; Gaps 4;

QY 7 LMSAGALALLA--PAFAOV--TPVTDELANPPAGEWISYGOENRHSPLTQITTT 61
 Db 16 LLSCAALAFSAVPAFAQEDTGTATSSDNGHP-GDWLSYSGSYEQRTSPDQINT 74
 QY 62 ENVGOLYVWANGMOPGKYOY-TPLIHDCVMT 92
 Db 75 ENVGKLKLMHYDLDTNRGQESTPLIVGVMT 106

RESULT 2
 DHET_ACEPO STANDARD: PRT: 738 AA.

ID DHET_ACEPO STANDARD: PRT: 738 AA.
 AC P28036: 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-NB11028;
 RX MEDLINE=91159482; Pubmed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horiuchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D00635; BAA00528.1; -;
 DR PIR: S14270; S14270.
 DR HSSP: Q9Z477; 1TFG.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT POTENTIAL.
 FT ALCOHOL DEHYDROGENASE [ACCEPTOR].

FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 738 AA: 80840 MW: 1E2B6ED7BCD92AF6 CRC64.

Query Match 26.4%; Score 131.5; DB 1; Length 738;
 Best Local Similarity 32.6%; Pred. No. 2.5e-06;
 Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

QY 10 SAGALALLAFAQVTPVTDE-----LLANPAGEWISYGOENRHSPLTQITTT 61
 Db 17 TAGTICAAALISGATWASADGATGTAIIHADHPENMMYTGRTTSDORTSPDQINR 76
 QY 62 ENVGOLYVWANGMOPGKYOY-TPLIHDCVMT 92
 Db 77 SNVGKLKLMHYDLDTNRGQESTPLIVGVMT 108

RESULT 3
 DHET_ACEEU STANDARD: PRT: 739 AA.

ID DHET_ACEEU STANDARD: PRT: 739 AA.
 AC Q44002: 007952;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconacetobacter.
 OX NCBI_TaxID=33995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DES1 / DSM 6160;
 RA Thurner C.A.K.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X82894; CAA58066.1; -;
 DR EMBL: Y09480; CAA70688.1; -;
 DR HSSP: Q9Z477; 1TFG.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 739
 FT BINDING 651 651 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 654 654 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 655 655 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 739 AA: 80944 MW: E681BB237ACB91F4 CRC64.

RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of Escherichia coli modelled on that of methanol dehydrogenase from Methylobacterium extorquens.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: PO4.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51323; CA35706.1; -
 DR EMBL: D12651; BA02174.1; -
 DR EMBL: D26562; CA820298.1; -
 DR EMBL: AE000122; AAC73235.1; -
 DR PIR: J01017; J01017.
 DR HSSP: P38539; 4AAH.
 DR Ecogene: EG10369; qcd.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR Oxidoreductase; POQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10
 FT TRANSMEM 11 37 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT ACT_SITE 142 796 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TMSADTP -> HIKRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9BDC705A12894E9 CRC64;

Query Match 25.6%; Score 127.5; DB 1; Length 796;
 Best Local Similarity 33.7%; Pred. No. 7,1e-06;
 Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LMAS-----AGATALLAFAAQTVPYVTELLANPPAGMISYQNOENRHSPLTQ 58
 DB 135 LTVAGNDPOEINGTSLSDATPAEA-ISPVAQ-----DWPAYKNGQGRSPKQ 185
 OY 59 ITTENYGOLOIYWA-----RGMOPGRV--QVTPLIHDGVYLT 93
 DB 186 INADNHNLKEAVERTGDKOPNDPGEITNEVTPIKGVDLYL 229

RESULT 6

DHG_GLUOX
 ID DHG_GLUOX STANDARD: PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OX NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92017653; PubMed-1833618;
 RA Cleon-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: PO4.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN HERE.
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; OPEX.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 KW SIGNAL
 FT SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54 POTENTIAL.
 FT TRANSMEM 59 76 POTENTIAL.
 FT TRANSMEM 94 108 POTENTIAL.
 FT TRANSMEM 123 138 POTENTIAL.
 FT ACT_SITE 470 470 BASE (POTENTIAL).
 FT ACT_SITE 788 788 H -> N (IN P2 FORM).
 FT VARIANT 788
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match 21.7%; Score 108; DB 1; Length 808;
 Best Local Similarity 28.2%; Pred. No. 0.00084;
 Matches 29; Conservative 14; Mismatches 40; Indels 20; Gaps 3;

OY 13 ALATLAAFAAQTVPYVTELLA-----NPPAGEWISYQNOENRHSPLTQ 61
 DB 128 AVLAFASLFTDPHISGLPTQIANSPADPNVPASEWAIYRTQAGDRWSPLAN 187
 OY 62 ENVGOLOIYWA-----ARGMOPGRV--QVTPLIHDGVYLT 95
 DB 188 TNSNLKAVMHIHTRKDMMSNDPGEOTNEATPIEFNNITLYWCS 230

RESULT 7
DHML_METOR STANDARD: PRT: 626 AA.
ID DHML_METOR
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large alpha subunit) (MEDH).
GN MOX.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RC STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
RX MEDLINE=89008094; PubMed=2459109;
RA Machlin S.M., Hanson R.S.;
RT "Nucleotide sequence and transcriptional start site of the methylobacterium organophilum XX methanol dehydrogenase structural gene";
RT J. Bacteriol. 170:4739-4747(1988).
RL
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).
CC -----
DR EMBL: M2629; AAA50289.1; -.
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; BACTERIAL_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT SIGNAL 1 28
FT CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 130 131 BY SIMILARITY.
FT DISULFID 413 442 BY SIMILARITY.
FT ACT_SITE 330 330 BASE (POTENTIAL).
SQ SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DF CRC64;
Query Match 20.9%; Score 104; DB 1; Length 626;
Best Local Similarity 34.8%; Pred. No. 0.0017;
Matches 31; Conservative 18; Mismatches 34; Indels 6; Gaps 5;
OY 10 SAGALALA-AP-AFAQVTPYTDLLANPAGE-WISTGQNEVTRSPLOTTEVVGQ 66
DB 7 SVSALAMLALAPALSSVAYANDKLEVELSDNMVPMKKNYSNNYSELKQVKNKSVKQ 66
OY 67 LQLVW--ARGMPKRVQVYTPPLIHGDMVL 93
DB 67 LRPAMTFSTGLNGH-EGAPLVVDGKMTY 94
RESULT 8
DHGA_ACICA STANDARD: PRT: 801 AA.
ID DHGA_ACICA

AC P05465;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase A [pyrroloquinoline-quinone] precursor (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDH-A.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMD 79.41;
RX MEDLINE=88289368; PubMed=339393;
RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
RT "Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from Acinetobacter calcoaceticus";
RL Nucleic Acids Res. 16:6228-6228(1988).
CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucose-1,5-lactone + reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERS 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).
CC -----
DR EMBL: X07235; CA30222.1; -.
DR PIR: S00943; S00943.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Transmembrane; Periplasmic; Signal.
FT SIGNAL 1 33
FT CHAIN 34 801
FT TRANSMEM 39 55 GLUCOSE DEHYDROGENASE-A
FT TRANSMEM 59 79 [PYRROLOQUINOLINE-QUINONE].
FT TRANSMEM 94 108 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT ACT_SITE 471 471 BASE (POTENTIAL).
SQ SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
Query Match 19.1%; Score 95; DB 1; Length 801;
Best Local Similarity 31.2%; Pred. No. 0.02;
Matches 24; Conservative 11; Mismatches 22; Indels 20; Gaps 3;
OY 20 PAFACVTPYTDLLANPAGEWISTGQNEVTRSPLOTTEVVGQLOLVNARGMPGK 79
DB 152 PETAAVPGVAB-----SDMPAYGRTOAGVRYSPKQINDQVKKVAMT--LRTGD 202
OY 80 V-----QVPL 85
DB 203 LKTNDGSETTNOVTP 219
RESULT 9
DHML_METEX STANDARD: PRT: 626 AA.
ID DHML_METEX

AC P16027;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MEDH).
 GN MOXF.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=90337342; PubMed=2116368;
 RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
 RT "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
 RL Gene 90:173-176(1990)."
 RN [2]
 RP SEQUENCE OF 28-53.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=89350892; PubMed=2504152;
 RA Nunn D.N., Day D., Anthony C.;
 RT "The second subunit of methanol dehydrogenase of Methylobacterium
 RL extorquens AM1".
 RN Biochem. J. 260:857-862(1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=95384759; PubMed=7656012;
 RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
 RT "The active site of methanol dehydrogenase contains a disulphide
 RL bridge between adjacent cysteine residues.";
 RN Nat. Struct. Biol. 1:102-105(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
 RX MEDLINE=95253818; PubMed=7735834;
 RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
 RT "The refined structure of the quinoprotein methanol dehydrogenase
 RL from Methylobacterium extorquens at 1.94 A.";
 RN Structure 3:177-187(1995).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M31108; AAA25380.1; -
 DR DR PIR: S07908; S07908.
 DR DR PIR: J00706; J00706.
 DR HSSP: P38539; AAAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 27
 FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330 BASE (POTENTIAL).
 FT

SO SEQUENCE 626 AA; 68434 MW; 64988D0AFD2AD34C CRC64;
 Query Match 18.9%; Score 94; DB 1; Length 626;
 Best Local Similarity 34.8%; Pred. No. 0.019;
 Matches 31; Conservative 16; Mismatches 36; Indels 6; Gaps 5;
 QY 10 SAGALLALA-AP-AFAQVPTVDLLANPAGE-WISYQONQENVRRSPQTTEWVQ 66
 Db 7 SVSALLMALAPALSSGAYANDKIVELSKSDNNVMGKNYDSNLSKQINKGVAK 66
 QY 67 LQLVW--ARGMPGKVQVTPPLIHGVMFL 93
 Db 67 LRPAMFTSTGLNGH-EGAPLVVDKMYI 94
 RESULT 10
 ID DHM1_PARDE STANDARD; PRT; 631 AA.
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MEDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 33-49.
 RX MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RL structural gene from Paracoccus denitrificans.";
 RN J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 CC TOTAL CELL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M17339; AAA8366.1; -
 DR DR HSSP: P38539; AAAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631
 FT DISULFID 135 136 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 448 447 BY SIMILARITY.
 FT ACT_SITE 335 335 BASE (POTENTIAL).
 FT SEQUENCE 631 AA; 69799 MW; 0934DC93PFC5730B CRC64;
 Query Match 18.7%; Score 93; DB 1; Length 631;
 Best Local Similarity 35.8%; Pred. No. 0.025;
 Matches 34; Conservative 15; Mismatches 34; Indels 12; Gaps 6;

RESULT	13		
PSAB_ODOST			
ID	PSAB_ODOST	STANDARD:	PRT, 733 AA.
AC	P9460;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PST-B).		
GN	PSAB.		
OS	Odontella sinensis.		
OG	Chloroplast.		
OC	Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;		
OC	Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.		
OX	NB01_TaxID=2839;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;		

RESULT	14	
PSAB_HEATTR		
ID	PSAB_HEATTR	STANDARD;
AC	Q9XCV2;	PRT; 776 AA.
DT	01-MAR-2002 (Rel. 41, Created)	
DT	01-MAR-2002 (Rel. 41, Last sequence update)	

01-MAR-2002 (Rel. 41, last annotation update)
 DE Photosystem I P700 chlorophyll A apoprotein A2 (psab) (PSI-B).
 GN PSAB.
 OS Heterocapsa triquetra (Dinoflagellate).
 OC Chloroplast.
 CC Eukaryota; Alveolata; Dinophyceae; Peridinales; Heterocapsaceae;
 CC Heterocapsa.
 CC NCBI_TaxID=66468;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-CMP 449;
 CC MEDLINE=99334925; PubMed=10408440;
 CC Zhang Z., Green B.R., Cavalier-Smith T.;
 CC "Single gene circles in dinoflagellate chloroplast genomes.";
 CC Nature 400:155-159(1999).
 CC -1- FUNCTION: Psab and psab bind P700, the primary electron donor of
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
 CC FX. Psi functions as a plastocyanin/cytochrome c6-ferredoxin
 CC oxidoreductase.
 CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
 CC -1- SUBUNIT: A psab/B heterodimer binds the P700 chlorophyll special
 CC pair and subsequent electron acceptors. The psi reaction center of
 CC higher plants and algae is composed of one at least 11 subunits.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF130032; AAD4699.1; -;
 CC InterPro: IPR001280; Psab_Psab.
 CC Pfam: PF00223; psab; psab; 2.
 CC PRINTS: PRO00257; PHOTOSYSAB.
 CC PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB: 1.
 CC Photosynthesis; Photosystem I; Electron transport; Chloroplast;
 CC Thylakoid; Transmembrane; Iron-sulfur 4Fe-4S; Chlorophyll.
 CC KW Thylakoid; Transmembrane; Iron-sulfur 4Fe-4S; Chlorophyll.
 CC FT TRANSMEM 57 80
 CC FT TRANSMEM 150 172
 CC FT TRANSMEM 233 257
 CC FT TRANSMEM 334 352
 CC FT TRANSMEM 382 405
 CC FT TRANSMEM 421 447
 CC FT TRANSMEM 472 494
 CC FT TRANSMEM 561 579
 CC FT TRANSMEM 619 640
 CC FT TRANSMEM 687 709
 CC FT TRANSMEM 753 773
 CC FT METAL 603 612
 CC FT METAL 612 612
 CC FT BINDING 698 698
 CC FT BINDING 706 706
 CC FT BINDING 714 714
 CC FT BINDING 715 715
 CC FT BINDING 746 746
 CC SEQUENCE 776 AA; 86143 MW; F8640CA5F58C8913 CRC64;

Query Match 14.6%; Score 72.5; DB 1; Length 776;
 Best Local Similarity 25.9%; Pred. No. 4.6;
 Matches 28; Conservative 18; Mismatches 39; Indels 23; Gaps 6;

QY 4 TSLMASAGALALALAPA---FAQVTPVTDELIANP-----AGEWISYGCNENT 51
 DB 244 TSLAM--AGHLIHAIPASRGIVHTISPVGDSFAPSLKALVPFVSGNMYAODIDKD 301

QY 52 RASPLQITENTNGQLQWARGMQ--PGKQVTPPLIHD-----GWATL 93
 DB 302 NH---IFGSPVAGAKALITFLPGVKSDPTASLYLTDIAHHHLAGVLFI 346
 RESULT 15
 ID XOXF_PARDE STANDARD; PRT; 600 AA.
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, last annotation update)
 DT 15-DEC-1998 (Rel. 37, last annotation update)
 DE Putative dehydrogenase XOXF precursor (Ec 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Paracoccus.
 CC NCBI_TaxID=266;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RA Harms N.;
 CC Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC [2]
 CC SEQUENCE OF 380-600 FROM N.A.
 CC STRAIN-PD 1235;
 CC MEDLINE=92041583; PubMed=1657873;
 CC Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
 CC Stoutamer A.H.;
 CC "Isolation, sequencing, and mutagenesis of the gene encoding
 CC cytochrome c551 of Paracoccus denitrificans and characterization of
 CC the mutant strain";
 CC J. Bacteriol. 173:6971-6979(1991).
 CC -1- COFACTOR: PQQ (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U34346; AAC4455.1; -;
 CC EMBL: M75583; AAA2574.1; -;
 CC PIR: A41378; A41378.
 CC DR HSSP: P38539; 4AAH.
 CC DR InterPro: IPR002372; Bac_PQQ_repeat.
 CC DR Pfam: PF01011; Bacterial_PQQ; 7.
 CC Oxidoreductase; PQQ; Signal.
 CC FT SIGNAL 1 21
 CC FT CHAIN 22 600
 CC FT ACT_SITE 318 318
 CC SEQUENCE 600 AA; 65159 MW; DCA996F1BC3A3CE CRC64;

Query Match 14.2%; Score 70.5; DB 1; Length 600;
 Best Local Similarity 29.0%; Pred. No. 5.7;
 Matches 27; Conservative 14; Mismatches 33; Indels 19; Gaps 4;

QY 10 SAGALALALAPAFAQVTPVTDELIANP-----PAGEWISYGCNENTRHSPLQITTE 62
 DB 6 NGACALALALMSGTA-----LANEGRAGRDROAPQWALQMGDYANTRYSTLDQINKD 56
 QY 63 NVGQLQLVW--ARGMQPGKQVTPPLIHQVYML 93
 DB 57 NVKDLRVAMTFSTGVLRGH--EGSPVIGDMYV 88

Search completed: May 24, 2002, 10:15:47
 Job time: 765 sec

Fri May 24 11:27:23 2002

us-08-934-506a-5_copy_1_95.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:07 ; Search time 93.93 Seconds
(without alignments)
174.966 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPTSLMASGALALLAAP.....QPKVQVTPPLIHGVMYLAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMEL_19:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_human:*
6: SP_invertebrate:*
7: SP_mammal:*
8: SP_mhc:*
9: SP_organelle:*
10: SP_plant:*
11: SP_todent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rvirus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	157.5	31.6	608	2	Q93RE9	Q93RE9 pseudogluc
2	139.5	28.0	742	2	Q53362	Q53362 acetobacter
3	130.5	26.2	691	2	Q9AF95	Q9AF95 pseudomonas
4	130	26.1	786	2	P95466	P95466 pantoea cit
5	128	25.7	695	2	Q9P9U2	Q9P9U2 pseudomonas
6	127	25.5	623	2	Q9AGW3	Q9AGW3 pseudomonas
7	121	24.3	785	16	Q9KRF6	Q9KRF6 rhizobium 1
8	118.5	23.8	708	2	Q46444	Q46444 comonomas t
9	113.5	22.8	777	16	Q92R83	Q92R83 rhizobium m
10	108.5	21.8	803	16	Q91115	Q91115 pseudomonas
11	107.5	21.6	698	2	Q9KH03	Q9KH03 alcaligenes
12	103	20.7	629	2	Q9A048	Q9A048 methylolact
13	99	19.9	601	2	Q9EYR8	Q9EYR8 rhizobium m
14	99	19.9	601	16	Q92MY9	Q92MY9 rhizobium m
15	92	18.5	633	2	Q24759	Q24759 hypomicrob
16	90	18.1	790	2	Q9X255	Q9X255 pantoea cit

17	87.5	17.6	601	2	P71509	P71509 methylolact
18	83.5	16.8	695	2	Q934G0	Q934G0 pseudomonas
19	73.5	14.8	1229	5	Q9VM05	Q9VM05 drosophila
20	72.5	14.6	306	16	Q9RY72	Q9RY72 deinococcus
21	70	14.1	407	16	Q9A5E2	Q9A5E2 caulobacter
22	70	14.1	877	4	Q9H3Q6	Q9H3Q6 homo sapien
23	70	14.1	878	4	Q9H3Q7	Q9H3Q7 homo sapien
24	70	14.1	957	4	Q14651	Q14651 homo sapien
25	69.5	14.0	1217	4	Q9UKW9	Q9UKW9 homo sapien
26	69.5	14.0	1234	5	Q24690	Q24690 drosophila
27	69	13.9	711	5	Q9YGD4	Q9YGD4 drosophila
28	68.5	13.8	447	16	Q9RU43	Q9RU43 deinococcus
29	68	13.7	223	3	Q99015	Q99015 trichoderma
30	68	13.7	599	2	Q91935	Q91935 methyllovora
31	68	13.7	856	16	Q9A6E5	Q9A6E5 caulobacter
32	68	13.7	878	4	Q9G2Z2	Q9G2Z2 homo sapien
33	68	13.7	901	4	Q9H195	Q9H195 homo sapien
34	68	13.7	1057	11	Q9EON4	Q9EON4 mus musculus
35	67.5	13.6	1085	5	Q24363	Q24363 drosophila
36	67	13.5	197	16	Q9A829	Q9A829 caulobacter
37	67	13.5	738	2	Q70031	Q70031 streptomyces
38	66.5	13.4	428	2	Q9LH85	Q9LH85 thiolactillu
39	66.5	13.4	647	16	Q9HUB1	Q9HUB1 pseudomonas
40	66.5	13.4	1516	16	Q9HVT2	Q9HVT2 pseudomonas
41	66	13.3	465	5	Q18914	Q18914 caenorhabdi
42	66	13.3	1017	5	Q9XVY2	Q9XVY2 lycantria d
43	66	13.3	1419	5	Q9VY13	Q9VY13 drosophila
44	66	13.3	1668	5	Q76930	Q76930 drosophila
45	65.5	13.2	1320	11	Q9UK25	Q9UK25 rattus norv

ALIGNMENTS

RESULT 1
ID Q93RE9 PRELIMINARY; PRT: 608 AA.
AC Q93RE9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucanobacter saccharotogenes.
OC Bacteria; Pseudoglucanobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shibata T., Saito Y.;
RT "Alcohol dehydrogenase."
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB046580; BAB62258.1;
SQ SEQUENCE 608 AA; 65101 MW; 0ACEC97AE11BA570 CRC64;

Query Match	31.6%	Score: 157.5;	DB 2;	Length 608;
Best Local Similarity	34.3%	Pred. No. 1.1e-08;		
Matches	36;	Conservative	15;	Mismatches 35;
				Indels 19;
				Gaps 1;
QY	10	SAGALALLAAPFAO	-----VTPYDELLANPPAEEMISYGOEN	50
DB	16	STALIASLSPAFQHDNNAAEPSKAGSAIENFQVYADDLAKNPANWPLIRGNTYG	75	
QY	51	YRHSPLTOITTEENGOLQVYARGMOPKQVTPPLIHGVMYLAN	95	
DB	76	WGSPLDQINKDNVGLQVWRTMEPGSNGEALAIYNGVIFLGN	120	
RESULT 2				
ID Q53362		PRELIMINARY;	PRT: 742 AA.	
AC Q53362;		Q44159;		

DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE
 OS Acetobacter pasteurianus (Acetobacter turbidans)
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 NCBI_TaxID=438;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=NC11380;
 RX MEDLINE=94042848; PubMed=8226628;
 RA Takemura H., Konô K., Horinouchi S., Beppu T.;
 RT "Production by ethanol of alcohol dehydrogenase activity in Acetobacter
 pasteurianus."
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL; D13893; BAA0252.1;
 DR HSP; Q924J7; 1FLG.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 28.0%; Score 139.5; DB 2; Length 742;
 Best Local Similarity 40.0%; Pred. No. 1.2e-06;
 Matches 34; Conservative 10; Mismatches 38; Indels 3; Gaps 2;

QY 11 AGALALAAPAFQVTVTEDEL-ANPAGEMISYSGONENRHPLOITTEVNGOLQ 68
 DB 23 AALPYAAVPAADGGQNTGEALIHDDHENVMLSTGRYSRSPLOQINRSVGLK 82
 QY 69 LVNARGQPKVQ-VPLINDGVY 92
 DB 83 LAMTYITDNRNGEATPVLVDGIMY 107

RESULT 3
 ID 09AF95 PRELIMINARY; PRT; 691 AA.
 AC 09AF95;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 NCBI_TaxID=86174;
 RN
 RP SEQUENCE FROM N.A.
 RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT "Characterization of the expression of two distinct alcohol
 dehydrogenases involved in butane metabolism in Pseudomonas
 butanovora."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF355798; AAK27220.2;
 SQ SEQUENCE 691 AA; 75070 MW; 4FC7ED20CDA14P64 CRC64;

Query Match 26.2%; Score 130.5; DB 2; Length 691;
 Best Local Similarity 42.9%; Pred. No. 1e-05;
 Matches 24; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 39 GEMTISYSGONENRHPLOITTEVNGOLQVAVRGMOPGK-VQVPLINDGVY 93
 DB 41 GEMTHTGYDAGTRYSLAQTDPNAKELGLVMSYDLESSRVEATPIVDSVMY 96

RESULT 4
 P95466

ID P95466 PRELIMINARY; PRT; 786 AA.
 AC P95466;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GLUCOSE DEHYDROGENASE [PYROLOQUINOLINE-OUTINONE] (EC 1.1.99.17).
 GN GDH.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 NCBI_TaxID=53336;
 RN
 RP SEQUENCE FROM N.A. AND ACTIVITY.
 RC STRAIN=1056R;
 RX MEDLINE=97133947; PubMed=8979341;
 RA Cha J.-S., Pujol C., Kado C.I.;
 RT "Identification and characterization of a Pantoea citrea gene encoding
 glucose dehydrogenase that is essential for causing pink disease of
 pineapple."
 RL Appl. Environ. Microbiol. 63:71-76(1997).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR -> D-GLUCONO-1,5-LACTONE +
 REDUCED ACCEPTOR.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
 CC -1- MISCELLANEOUS: RESPONSIBLE FOR PINK TO BROWN DISCOLORATIONS OF
 INFECTED PINEAPPLES UPON HEATING DURING THE CANNING PROCEDURE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 DR EMBL; X95985; CAA65229.1;
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase; PQQ; Transmembrane; Cytoplasmic; Periplasmic.
 FT DOMAIN 1 13
 FT TRANSMEM 14 55
 FT DOMAIN 56 59
 FT TRANSMEM 60 80
 FT DOMAIN 81 85
 FT TRANSMEM 86 106
 FT DOMAIN 107 115
 FT TRANSMEM 116 136
 FT DOMAIN 137 431
 FT TRANSMEM 432 452
 FT DOMAIN 453 652
 FT TRANSMEM 653 673
 FT DOMAIN 674 700
 FT TRANSMEM 701 721
 FT DOMAIN 722 786
 FT ACT_SITE 724 770
 FT DOMAIN 770 786
 FT ACT_SITE 786 786
 SQ SEQUENCE 786 AA; 86038 MW; 372402AAD8067CC CRC64;

Query Match 26.1%; Score 130; DB 2; Length 786;
 Best Local Similarity 34.2%; Pred. No. 1.4e-05;
 Matches 35; Conservative 14; Mismatches 35; Indels 26; Gaps 5;

QY 2 KPTSLWASAGALALAAPAFQ-----VPTVTEDELANPP-----AGEMISYSGONQ 48
 DB 115 KPLAWYIGGINALLLGASLNDPOEINGVLNWSK-----PPASASASADWPAYGRTO 170
 QY 49 ENRHASPLTQITTEVNGOLQV-----ARGMOPGK-VQVPLINDGVY 93
 DB 171 EGVRYSPLTQINDKNVQOLQVAMQFHTGDHKTANDPGETINVTPLKGNMILY 224

RESULT 5
 ID 09F9U2 PRELIMINARY; PRT; 695 AA.
 AC 09F9U2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Bralnard J.B.,
 RA Terwilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2."
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF176640; AAC09249.1;
 DR HSSP; Q9Z457; 1F1G.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR InterPro; IPR000345; CytC_heme_bind.
 DR InterPro; IPR003088; CytC_C1.
 DR Pfam; PF01011; Bacterial_PQO; 6.
 DR Pfam; PF00034; Cytochrome_C; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 25.7%; Score 128; DB 2; Length 695;
 Best Local Similarity 38.3%; Pred. No. 1.9e-05;
 Matches 31; Conservative 14; Mismatches 32; Indels 4; Gaps 3;
 OY 15 ALAAPFAQVTPVTDLL--ANPPAGWISYQNGQENYRHSPLQITTEVNGQLQVMA 72
 ID 09AGW3 PRELIMINARY; PRT; 623 AA.
 AC 09AGW3;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sayavedra-Soto L.A., Vangnai A.S., Aip D.J.;
 RA "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF326086; AAK15506.1;
 DR HSSP; Q9Z457; 1F1G.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 4.
 KW Signal.
 FT SIGNAL 1 29 POTENTIAL.
 SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 25.5%; Score 127; DB 2; Length 623;
 Best Local Similarity 38.6%; Pred. No. 2.2e-05;
 Matches 39; Conservative 12; Mismatches 34; Indels 16; Gaps 5;
 OY 2 KPTSLMASGALALAPAFQVTPVTDLLAN--PPAGWISYQNGQENYRHSPLQIT 59
 ID 09AGW3 PRELIMINARY; PRT; 623 AA.
 AC 09AGW3;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sayavedra-Soto L.A., Vangnai A.S., Aip D.J.;
 RA "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF326086; AAK15506.1;
 DR HSSP; Q9Z457; 1F1G.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 4.
 KW Signal.
 FT SIGNAL 1 29 POTENTIAL.
 SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

DB 7 KPFLRAIVATAAALIPA-AAVTDVWEDIANDKTKTGVLTYGLKAKORHSPKAI 65
 OY 60 TTEVNGQLQVMA-----RGMOPGKVOVTPPIHDGVY 92
 ID 098KF6 PRELIMINARY; PRT; 785 AA.
 AC 098KF6;
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN MLI1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFP303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP002997; BAB4858.1;
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 7.
 KW Complete proteome.
 SQ SEQUENCE 785 AA; 84083 MW; 1DC224B7A1C25470 CRC64;

Query Match 24.3%; Score 121; DB 16; Length 785;
 Best Local Similarity 31.5%; Pred. No. 0.00013;
 Matches 39; Conservative 10; Mismatches 35; Indels 40; Gaps 4;
 OY 8 WASAG-----ALALAPAFQVTPVTD-----ELLANPAGE 40
 ID 046444 PRELIMINARY; PRT; 708 AA.
 AC 046444;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE QUINOHAEOMPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 DE (EC 1.1.99.-) (OH-EDH1).
 GN OHEDH.
 OS Commamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Commamonadaceae; Commamonas.
 OX NCBI_TaxID=285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=96184549; PubMed=8654419;

RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 RA De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinohemoprotein ethanol
 RT dehydrogenase of Comamonas testosteroni.";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RP SEQUENCE OF 32-54 AND 477-490.
 RC STRAIN-ATCC 15667;
 RX MEDLINE-95324580; PubMed-7601151;
 RA De Jong G.A.H., Geerloff A., Stoorvogel J., Jongejan J.A., De Vries S.,
 RA Duine J.A.;
 RT "Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 RT Purification, characterization, and reconstruction of the apoenzyme
 RT with pyriroloquinoline quinone analogues.";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE-86242113; PubMed-3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;
 RA "Quinohemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 RT testosteroni.";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION.
 RX MEDLINE-21536088; PubMed-11679760;
 RA Oubrie A., Huizinga E.G., Roseboom H.J., Kalk K.H., de Jong G.A.H.,
 RA Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinohemoprotein alcohol dehydrogenase from
 RT Comamonas testosteroni: crystals with unique optical properties.";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POO, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -
 DR HSSP: Q924J7; FFLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYCHROME.C1.
 KW Signal; POO; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635 QUINOHAEOPROTEIN ETHANOL DEHYDROGENASE
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BD6ACCA3 CRC64;

Query Match 23.8%; Score 118.5; DB 2; Length 708;
 Best Local Similarity 30.5%; Pred. No. 0.00021;
 Matches 32; Conservative 18; Mismatches 40; Indels 15; Gaps 4;

OY 3 PSLMASAGALALAAFAFAVPTV-----DELANP-AGWISTYGOEN 50
 DB 12 PGRMWLAACLG--SAAPAFQGTGPAQAAAVGVVDGDFIRANMARTPMPITGVDAE 69
 OY 51 YRHSPLQITTEVNGQLQVWARGMPGK-VQVPLIHGVMYLA 94
 DB 70 TRYSLDQINANVKDGLANSYMLSTRGVATPVVDDIMYVS 114

RESULT 9
 ID Q92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYROLOQUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021.
 RX MEDLINE-21368234; PubMed-11474104;
 RA Galibert F., Flan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Botte G.,
 RA Bouly M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
 RA Covle A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
 RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 RA Ramsperger U., Surzycki R., Thebaud P., Vandebol M.,
 RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
 RL Science 293:668-672(2001).
 DR EMBL: AL591785; CAC4553.1; -
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 777 AA; 84174 MW; D2F478A3BB10430 CRC64;

Query Match 22.8%; Score 113.5; DB 16; Length 777;
 Best Local Similarity 40.0%; Pred. No. 0.00082;
 Matches 28; Conservative 5; Mismatches 24; Indels 13; Gaps 2;

OY 35 NPAGEWISGQNGENYRHSPLQITTEVNGQLQVWARGMPGK-----QVT 83
 DB 159 NVPPGEHMOGRFPGQRYSPIDIRPENVANLQVAMQ--VQTDVYKLPEDVSETTYQVT 216
 OY 84 PLIHGVMYLA 93
 DB 217 PLKVKDTLYV 226

RESULT 10
 ID Q91115 PRELIMINARY; PRT; 803 AA.
 AC Q91115;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GCD OR PA2290.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004654; AAC05678.1; -
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO; 7.
 KW Complete proteome.

SEQUENCE 803 AA; 86217 MW; 7DDFA6B0FE5ED249 CRC64;

Query Match
Best Local Similarity 27.5%; Pred. No. 0.003; DB 16; Length 803;
Matches 36; Conservative 14; Mismatches 34; Indels 47; Gaps 5;

OY 6 LHM-----ASAGALALLAFAFAQVPTVDLLANP-----36
DB 103 LHMPLPFRRLADGPAFLGTAALGVAAYLAGAAVGSQ--FTNPGQIVGRIDRDSGMTST 161
OY 37 ----PAGEMISTGONENRHSPLQITTEENGOQLVWARMQSGV-----Q 81
DB 162 APAMPDGMQAYGRTEFFGDRYSPLQITPANVQLEAMR--IRTGDLPTADDPLELINE 219
OY 82 VVPLIDGVWY 92
DB 220 NTPKAVNGMIX 230

RESULT 11

O9KH03 PRELIMINARY; PRT; 698 AA.

AC O9KH03;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TERARYDROFUFURYL ALCOHOL DEHYDROGENASE.
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
ON NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2112557; PubMed=11222593;
RA Zarrat G., Schrader T., Andreessen J.R.,
RT "Catalytic and Molecular Properties of the Quinohemoprotein
Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo.";
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL; AF277373; AAF8635.1; -.
DR HSSP; Q924J7; 1FLG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR00345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt.CI.
DR Pfam; PF01011; Bacterial_PQQ_6.
DR Pfam; PF00034; Cytochrome_c_1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOMN_1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180DBID2FB2 CRC64;

Query Match 21.6%; Score 107.5; DB 2; Length 698;
Best Local Similarity 34.4%; Pred. No. 0.003; DB 2; Length 698;
Matches 31; Conservative 14; Mismatches 40; Indels 5; Gaps 3;

OY 10 SAGALALLAFAFA--GVTPVTDLLANPAG--EWISYSGONENRHSPLQITTEENVG 65
DB 14 AASATLPAFAFANAAARVDGAIRANRGNPSTGYAETRFSLKEOVNAGNVR 73
OY 66 QLQVWARMQSGV--VQVTPILIHGVWYLA 94
DB 74 NLGLAMSYDLESTGVEATPLVVGVMIVS 103

RESULT 12

O9A048 PRELIMINARY; PRT; 629 AA.

AC O9A048;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE MXAF.
GN MXAF.

OS Methylobacterium nodulans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Methylobacterium group; Methylobacterium.

ON NCBI_TaxID=114616;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=ORS2060;

RX MEDLINE=20566686; PubMed=11114919;

RA Sy A., Giraud E., Jourand P., Garcia N., Willems A., de LaJude P.,

RA Prin Y., Neyra M., Gillis M., Bolvin-Messon C., Dreyfus B.;

RT "Methylobacterium nodulans: a bacterium that fix nitrogen in

RT symbiosis with legumes.";

RL J. Bacteriol. 183:214-220(2001).

DR EMBL; AF220764; AAC49450.1; -.

DR HSSP; P38539; 4AAH.

DR InterPro: IPR001479; Bac_PQQ.

DR InterPro: IPR002372; Bac_PQQ_repeat.

DR Pfam; PF01011; Bacterial_PQQ_7.

DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.

DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.

SQ SEQUENCE 629 AA; 68798 MW; 3DB559D8373BB0B2 CRC64;

Query Match 20.7%; Score 103; DB 2; Length 629;
Best Local Similarity 32.3%; Pred. No. 0.0087;
Matches 30; Conservative 18; Mismatches 37; Indels 8; Gaps 4;

OY 4 TSLMASAGALALLAFAFAQVPTVDLL--ANPPAGEMISYSGONENRHSPLQITTE 62
DB 11 TGVSAALALLPLPLGFS-----ALANDKVLKSKSDGWMVPGKNYDSDNYSKLQINAE 66
OY 63 NVGQQLVW--ARGMQPGKVOVTPILIHGVWYL 93
DB 67 NVKNLKVSWQFSTGLNGH--EGAPLVVDGTMVY 98

RESULT 13

O9EYW8 PRELIMINARY; PRT; 601 AA.

AC O9EYW8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
GN MXAF.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM1021;
RA Fennel B.J., Tiwari R.P., Dilworth M.J.;
RT "Regulation of C1 assimilation in Sinorhizobium meliloti.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF309488; AAG31643.1; -.
DR HSSP; P38539; 4AAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam; PF01011; Bacterial_PQQ_7.
SQ SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 19.9%; Score 99; DB 2; Length 601;
Best Local Similarity 34.5%; Pred. No. 0.022;
Matches 29; Conservative 17; Mismatches 32; Indels 6; Gaps 4;

OY 14 LALLAFAFAQVPTVDLL--LANPPAGEMISYSGONENRHSPLQITTEENGOQLVW 71
DB 8 LAIMSTGGAGVAFANDELKLDLIDP--NOMAIOTGDYANLRYSKIDQINDNKGKLVAM 66
OY 72 --ARGMQPGKVOVTPILIHGVWYL 93
DB 67 TFSYGLRGH--EGSPLVIGDLMVY 89

RESULT 14
ID 092WY9 PRELIMINARY; PRT: 601 AA.
AC 092WY9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
GN SM20173.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Flan T.M., Weldner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoelster F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Goiding B., Puenler A.;
RT "The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL: AL603642; CAC48573.1;
KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 601 AA; 65759 MW; D73424EFED15ADBE CRC64;

Query Match 19.9%; Score 99; DB 16; Length 601;
Best Local Similarity 34.3%; Pred. No. 0.022;
Matches 29; Conservative 17; Mismatches 32; Indels 6; Gaps 4;

QY 14 LALLAFAFAQVPTVDEL--LANPAGMEISYQNOENYRHSPLQITTEHYGOLQVW 71
DB 8 LAIMSIGGGAQVAFANDELQKIDP--NQNAIQGTANLRLSKLDQINKDNQKQVAV 66
QY 72 --ARGMOPKQVPTPLHDCGYWL 93
DB 67 TFGTVLKGH-EGSPVIVGIDLMYV 89

RESULT 15
ID 024759 PRELIMINARY; PRT: 633 AA.
AC 024759;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN KXAF.
OS Hyphomicrobium methylovorum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=84;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM2;
RX MEDLINE=97457202; PubMed=9311140;
RA Tanaka Y., Yoshida T., Watanabe K., Izumi Y., Mitsunaga T.;
RT "Cloning and analysis of methanol oxidation genes in the methylovor
Hyphomicrobium methylovorum GM2.";
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL: AB004097; BAA23272.1;
DR HSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_P00.
DR InterPro: IPR002372; Bac_P00_repeat.
DR Pfam: PF01011; Bacterial_P00_7.
DR PROSITE: PS00363; BACTERIAL_P00_1.
DR PROSITE: PS00364; BACTERIAL_P00_2; 1.

KW Signal: Oxidoreductase.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 633 POTENTIAL.
SQ SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;

Query Match 18.5%; Score 92; DB 2; Length 633;
Best Local Similarity 28.3%; Pred. No. 0.14;
Matches 26; Conservative 20; Mismatches 40; Indels 6; Gaps 3;

QY 6 LMASAGALALLAFAFAQVPTVDELANPAG--WISYQNOENYRHSPLQITTEHY 64
DB 12 LMSASCAIATVALQVYASSAYANDKLIELSKNENWVPGKNYDSNNYSESTQVAAENV 71
QY 65 GOLQVYARGMOPKQV---QVPTPLHDCGYWL 93
DB 72 KOLKHAWS--FTSGELHGHGAPLVIGDMYV 101

Search completed: May 24, 2002, 10:17:28
Job time: 741 sec

5
Fri May 24 11:27:24 2002

us-08-934-506a-5_copy_1_95.rspt

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:04:59 ; Search time 117.28 Seconds
(without alignments)
118.385 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPRLMASGALALALAP.....KTGDLWEHRRLPNTATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A-Geneseq.032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution...

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	100.0	579	19	AAW37873 Alcohol and/or ald
2	587	89.3	579	19	AAW37874 Alcohol and/or ald
3	548	83.4	579	19	AAW37876 Alcohol and/or ald
4	534.5	81.4	578	19	AAW37875 Alcohol and/or ald
5	189.5	28.8	754	21	AAW35987 Sorbitol dehydrog
6	183	27.9	738	12	AAW35993 A.altoacetogenes me
7	177	26.9	738	13	AAW20192 ADH complex protei
8	150.5	22.9	742	11	AAW05235 Amino acid sequenc
9	145.5	22.1	740	22	AAW24430 Novel human diageno
10	129	19.6	948	22	AAW95019 Sorbitol dehydrog
11					Novel human diageno

12	129	19.6	1510	22	ABG21573
13	76	11.6	824	18	AAW09614
14	73.5	11.2	1229	22	ABW71736
15	72.5	11.0	343	19	AAW35001
16	72.5	11.0	462	17	AAW05395
17	71	10.8	526	22	AAW99895
18	70	10.7	1016	22	AAW34349
19	69.5	10.6	275	21	AAW42781
20	69.5	10.6	1419	22	ABW59208
21	69	10.5	711	22	ABW69743
22	69	10.5	957	22	ABW62401
23	69	10.5	1029	22	ABW6194
24	69	10.5	1029	22	ABW62658
25	68.5	10.4	1029	22	ABW1927
26	68	10.4	424	22	ABW59716
27	67.5	10.3	171	22	AAU32330
28	67.5	10.3	275	21	AAW16656
29	67.5	10.3	1041	22	ABG21594
30	67.5	10.3	1224	18	AAW6735
31	67	10.2	1224	18	AAW63034
32	67	10.2	159	19	AAW4325
33	67	10.2	974	19	AAW49079
34	67	10.2	1407	22	ABG28517
35	67	10.2	1429	22	ABW8779
36	66.5	10.1	333	22	ABW63676
37	66.5	10.1	384	22	ABW85251
38	66.5	10.1	551	22	ABW85252
39	66.5	10.1	551	22	ABW85252
40	66.5	10.1	577	22	ABW08719
41	66	10.0	227	22	AAU56221
42	66	10.0	620	21	AAW77208
43	66	10.0	658	22	ABW6214
44	66	10.0	673	22	ABW59674
45	65.5	10.0	271	18	AAW28260

ALIGNMENTS

RESULT	1	AAW37873	standard; Protein; 579 AA.
AAW37873			
XX			
AC	AAW37873;		
XX			
DT	10-AUG-1998	(first entry)	
XX			
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.		
XX			
KW	alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;		
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;		
KW	2-keto-L-gulonate; L-ascorbic; inhibition.		
XX			
OS	Gluconobacter oxydans.		
XX			
FT	Key	Location/Qualifiers	
FT	Peptide	1..23	
FT	Protein	/note="signal peptide"	
FT		24..579	
FT		/note="mature protein"	
PN	EP832974-A2.		
XX			
PD	01-APR-1998.		
XX			
PF	11-SEP-1997;	97EP-0115801.	
XX			
PR	19-SEP-1996;	96EP-0115001.	
XX			
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.		
XX			
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;		
XX			

Novel human diageno
Pathogenic Staphyl
Drosophila melanog
Bacillus thymoloe
Human SH3P17 prote
Xanthomonas sp. ch
Staphylococcus aur
Arabidopsis thalia
Drosophila melanog
Drosophila melanog
Drosophila melanog
Novel human diageno
Novel human diageno
Acromonium cellulo
Drosophila melanog
Novel human secret
Arabidopsis thalia
Novel human diageno
Staphylococcus car
Mycobacterium tube
Mycobacterium tube
Solium tuberosum
Novel human diageno
Drosophila melanog
Drosophila melanog
Plant porobilinog
Plant thioedoxin-
Novel human diageno
Propionibacterium
S. venezuelae deso
Drosophila melanog
Drosophila melanog
Amino acid sequenc

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 XX
 SQ Sequence 579 AA;

Query Match 100.0%; Score 657; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2,2e-64;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALIALAPAFQVTPVTDELLANPPAGEMISYQONENYRHSPLTQIT 60
 DB 1 MKPTSLIMASAGALIALAPAFQVTPVTDELLANPPAGEMISYQONENYRHSPLTQIT 60
 QY 61 TENVGQLQLVWARGMOPGKQVTPPLIHGVMYLANPGDVIAIDAKTDDLWEHRRLPN 120
 DB 61 TENVGQLQLVWARGMOPGKQVTPPLIHGVMYLANPGDVIAIDAKTDDLWEHRRLPN 120
 QY 121 IATLN 125
 DB 121 IATLN 125

RESULT 2
 AAW37874
 ID AAW37874 standard; Protein: 579 AA.
 AC AAW37874;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key
 FT Peptide 1..23 Location/Qualifiers
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 PN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 XX
 SQ Sequence 579 AA;

Query Match 89.3%; Score 587; DB 19; Length 579;
 Best Local Similarity 91.2%; Pred. No. 1.3e-56;
 Matches 114; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALIALAPAFQVTPVTDELLANPPAGEMISYQONENYRHSPLTQIT 60
 DB 1 MKPTSLIMASAGALIALAPAFQVTPVTDELLANPPAGEMISYQONENYRHSPLTQIT 60
 QY 61 TENVGQLQLVWARGMOPGKQVTPPLIHGVMYLANPGDVIAIDAKTDDLWEHRRLPN 120
 DB 61 TENVGQLQLVWARGMOPGKQVTPPLIHGVMYLANPGDVIAIDAKTDDLWEHRRLPN 120
 QY 121 IATLN 125
 DB 121 IATLN 125

RESULT 3
 AAW37876
 ID AAW37876 standard; Protein: 579 AA.
 AC AAW37876;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key
 FT Peptide 1..23 Location/Qualifiers
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 PN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR WPI; 1998-195228/18.
 DR N-PSDB; AAV29054.
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 44-46; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA;

Query Match 83.4%; Score 548; DB 19; Length 579;
 Best Local Similarity 81.6%; Pred. No. 2.9e-52;
 Matches 102; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

OY 1 MKPTSLMASAGALALLAAPAFQVTPVDELLANPPAGEWISTGQNDENRHSPLTQIT 60
 |||||
 Db 1 mptclltsaavllltapaafaycpicdelampagewinyngenythspitqit 60
 OY 61 TENVGQLOLVWARGMOPKQVOTPLIHDSVMYLANPGDVIAIDAKTGDLIWEHRROLPN 120
 :|||||
 Db 61 adhwgqqlvawrgmeagvqvtpmldhgvmylanpgdvigaldqgtglivewhrqlpn 120
 OY 121 IATLN 125
 :||||
 Db 121 vatln 125

RESULT 4
 AAW37875
 ID AAW37875 standard; Protein; 578 AA.
 XX
 AC AAW37875;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note- "signal peptide"
 FT 24..578
 FT /note- "mature protein"

EP832974-A2.
 01-APR-1998.
 11-SEP-1997; 97EP-0115801.
 19-SEP-1996; 96EP-0115001.
 (HOFF) HOFFMANN LA ROCHE & CO AG F.
 Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI; 1998-195228/18.
 DR N-PSDB; AAV29053.
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A'' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 578 AA;

Query Match 81.4%; Score 534.5; DB 19; Length 578;
 Best Local Similarity 81.6%; Pred. No. 9.1e-51;
 Matches 102; Conservative 9; Mismatches 13; Indels 1; Gaps 1;

OY 1 MKPTSLMASAGALALLAAPAFQVTPVDELLANPPAGEWISTGQNDENRHSPLTQIT 60
 |||||
 Db 1 mkltltlqssaallvltipalag-taltdemlanppagewinyngenythspitqit 59
 OY 61 TENVGQLOLVWARGMOPKQVOTPLIHDSVMYLANPGDVIAIDAKTGDLIWEHRROLPN 120
 :|||||
 Db 61 adhwgqqlvawrgmeagkigvcpvldhgvmylanpgdvigaldqgtglivewhrqlpn 119
 OY 121 IATLN 125
 :||||
 Db 120 iatln 124

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 XX
 AC AAB35987;
 XX
 DT 01-MAR-2001 (first entry)
 XX
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulononic acid.
 XX
 OS Gluconobacter oxydans.
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-1B00736.
 XX
 PR 23-APR-1999; 99WO-1B00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI; 2000-687351/67.
 DR N-PSDB; AAC83153.
 XX
 PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative

PF Production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol -
 XX Claim 1; Fig 8; 96pp; English.
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 XX Sequence 754 AA:
 SQ

Query Match 28.8%; Score 189.5; DB 21; Length 754;
 Best Local Similarity 39.0%; Pred. No. 3,1e-12;
 Matches 46; Conservative 17; Mismatches 46; Indels 9; Gaps 3;

OY 14 LALLAARFAQVTPV-----TDELLANPAGEMISYGONENYRHSPLQITTEENYGO 66
 Db 17 Lgcnaalacatspvalaedtgaltadqhpqdwmsygrtysqrppldtkdnasn 76
 OY 67 LQLVWARGMQPKQVQV-TPLIHGVMYLANPGDVIOAIDAKTGDLIMEHRQLP-NTA 122
 Db 77 Lklawhyldltnrggegrplivdygvmattmwmkmdaatgkllwsydkvpqna 134

RESULT 6

AAR13993
 ID AAR13993 standard; Protein; 738 AA.

AC AAR13993;

DT 09-DEC-1991 (first entry)

DE A. altocetigenes membrane-bound ADH 72kd sub-unit.

KW alcohol dehydrogenase complex; carboxylic acid production.

OS Acetobacter altocetigenes.

PN EP448969-A.

PD 02-OCT-1991.

PF 26-FEB-1991; 91EP-0102793.

PR 26-MAR-1990; 90JP-0073440.

PR 26-FEB-1990; 90JP-0042391.

PA (NAKA-) NAKANO VINEGAR KK.

PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;

DR WPI; 1991-289462/40.

DR N-PSDB; AAQ13580.

PT Gene for membrane-bound alcohol dehydrogenase complex - obtd.
 from Acetobacter altocetigenes, used for prodn. of enzyme for
 converting alcohol to acid

PS Disclosure; Fig 3; 36pp; English.

CC Total DNA was prepared from A. altocetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation

CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. altocetigenes (see AAQ13582-013584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 CC
 XX Sequence 738 AA:
 SQ

Query Match 27.9%; Score 183; DB 12; Length 738;
 Best Local Similarity 33.3%; Pred. No. 1,6e-11;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGALALLAARFAQVTPVDE-----LLANPAGEMISYGONENYRHSPLQITTT 61
 Db 17 tagtcaallsgyatmasaddggagatgaalhadddhpnwmtygrtysdqryspldqlnr 76
 OY 62 ENVGQLVWARGMQPKQVQV-TPLIHGVMYLANPGDVIOAIDAKTGDLIMEHRQLP- 119
 Db 77 snvgnlklawylldltnrggegrplivdygvmattmwmkmdaatgkllwsydkvpq 136
 OY 120 NTA 122
 Db 137 nla 139

RESULT 7

AAR20192
 ID AAR20192 standard; Protein; 738 AA.

AC AAR20192;

DT 16-APR-1992 (first entry)

DE ADH complex protein (mol.wt. 72.000).

KW Alcohol dehydrogenase; acetic acid; fermentation.

OS Acetobacter altocetigenes NH-24.

PN JP03266988-A.

PD 27-NOV-1991.

PF 26-MAR-1990; 90JP-0073440.

PR 26-FEB-1990; 90JP-0042301.

PR 26-MAR-1990; 90JP-0073440.

PA (NAKA-) NAKANO SUTEN KK.

PI WPI; 1992-019325/03.

DR N-PSDB; AAQ20383.

PT Alcohol dehydrogenase complex structural gene - used in plasmid
 and enhancing efficiency of acetic acid fermentation for
 transformed acetic acid bacteria

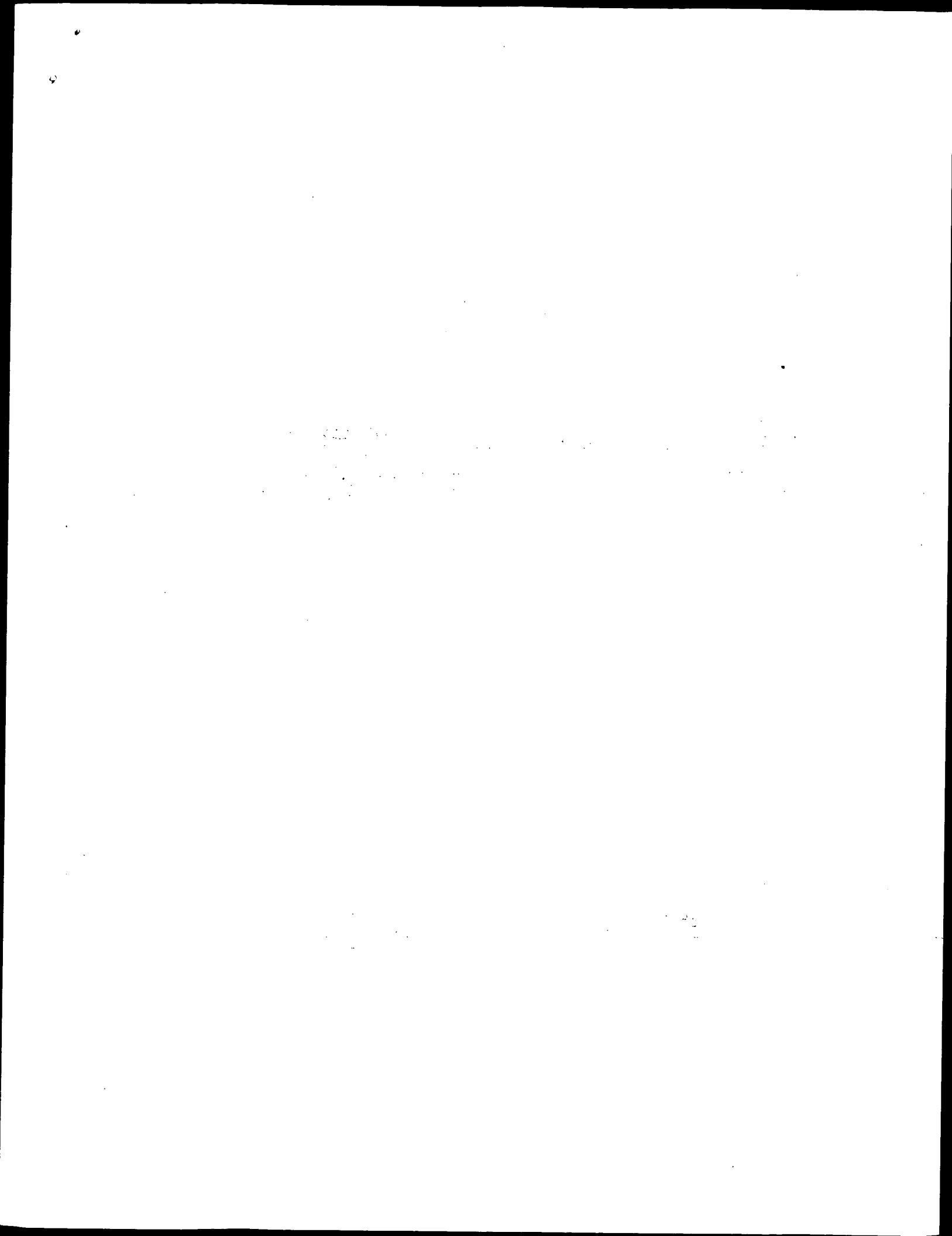
PS Disclosure; Fig 3(1-3); 21pp; Japanese.

CC Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex
 CC can be easily extracted from the bacteria and purified and it can be
 CC used for the determination of an alcohol.
 CC See also AAQ20383-84, and -86-88.

CC Sequence 738 AA;

Fri May 24 11:27:12 2002

us-08-934-506a-5_copy_1_125.rag



RESULT 2
US-07-985-458-3
Sequence 3, Application US/07985458
Patent No. 5344777
GENERAL INFORMATION:
APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takaya, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshuya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
TITLE OF INVENTION: Bacteria

```

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frisshauf, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: Purified protein having a molecular weight of about
IDENTIFICATION METHOD: 72,000
ORIGINAL SOURCE:
ORGANISM: Acetobacter alioacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tanaka, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Tayama, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshio;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horiuchi, Suenaru and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

```

```

Query Match          27.9%; Score 183; DB 1; Length 738;
Best Local Similarity 33.3%; Pred. No. 6.1e-13;
Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;
QY 10 SAGALALLAAPFAQVTPPTDE-----LLANPPAGEMTSYGNQENYRHSPLTQTTT 61
DB 17 TAGTICALISGYATMASDGGATGATTAIIHADDPGNMWTGRTYSQRYSPLDQINR 76

```

```

QY 62 ENVGQLQVWARGMQPGKQVY-TPLIHDGVYLANPGDVIQDAIDAKTGDLIWEHRQLP-119
DB 77 SNVGNLKLAWYLDLDTNRNGQGEFTPLVDGVYATTNMNMKAVDAATGKLMSYDPRVPG 136
QY 120 NIA 122
DB 137 NIA 139
RESULT
US-09-296-284-25
Sequence 25 Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Glucosobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533.087000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Glucosobacter suboxydans
US-09-296-284-25

```

```

Query Match          27.5%; Score 180.5; DB 4; Length 720;
Best Local Similarity 40.4%; Pred. No. 1.2e-12;
Matches 40; Conservative 18; Mismatches 36; Indels 5; Gaps 3;
QY 26 TPVTELLANPPAGEWISYGNQENYRHSPLTQITTEVNGQLVWARGMQPGKQVY-TP 84
DB 5 FAITN---ADQHPGDMMSGRYTSFORYSPLDQITKDNASNLKLMWYDLDITNRQGEFTP 61
QY 85 LHDGVMTYLANPGDVIQDAIDAKTGDLIWEHRQLP-NIA 122
DB 62 LIVDGVMTYTNMNMKALDAATGKLMSYDPKVGANIA 100

```

```

RESULT
US-09-136-251-2
Sequence 2, Application US/09136251A
Patent No. 6127156
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: OJIMA, Setsuko
APPLICANT: SHINJOH, Masako
TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
CURRENT APPLICATION NUMBER: US/09/136,251A
CURRENT FILING DATE: 1998-08-19
EARLIER APPLICATION NUMBER: EP 97114432.4
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 739
TYPE: PRT
ORGANISM: Glucosobacter suboxydans
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(24)
US-09-136-251-2

```



```

1 ADDRESS: Fish & Richardson, P.C.
2 STREET: 45 Rockefeller Plaza, Suite 2800
3 City: New York
4 STATE: NY
5 COUNTRY: US
6 ZIP: 10111
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette
10 COMPUTER: IBM compatible
11 OPERATING SYSTEM: Windows95
12 SOFTWARE: FASTSQ for Windows Version 2.0.
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/796,792
15 FILING DATE: 06-FEB-1997
16 CLASSIFICATION: 536
17
18 PRIORITY APPLICATION DATA:
19 APPLICATION NUMBER: 60/011,364
20 FILING DATE: 09-FEB-1996
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Hone, William J.
23 REGISTRATION NUMBER: 26,739
24 REFERENCE/DOCKET NUMBER: 07763/03301
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 212-765-5070
27 TELEFAX: 212-258-4291
28 INFORMATION FOR SEQ ID NO: 2:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 159 amino acids
31 TYPE: amino acid
32 TOPOLOGY: linear
33 MOLECULE TYPE: protein
34 FRAGMENT TYPE: internal
35
36 JS-08-796-792-2

```

Query Match 10.2%; Score 67; DB 3; Length 159;
 Best Local Similarity 21.8%; Pred. No. 2.4;
 Matches 36; Conservative 18; Mismatches 49; Indels 62; Gaps 7;

QY 1 MKPISLWASGALALAAPFAQVPTVDLLANPPAGWISYGOQENYRHSPLTQIT 60
 Db 6 MKTAVAVVAAAIATPAAPALAPPTGKL-----GSELTM-- 43

QY 61 TENYGOLOLVW-----ARGMOPG-----KVQVTPLIH-----DG 89
 Db 44 TDYVGQVVLGKVKSDKSTAVIGYIPVAGQWMEATATVNAIRGSYTPAVSQFNATRADG 103

QY 90 WY-----LANP-----GDVQALDAKTGDLIWEHRLPNIATLN 125
 Db 104 INRYVLMQAGPDRISGATIFQGEQSTGKIXFYDVYGPSPTIVAMN 148

RESULT 10
 US-08-868-786-6
 ; Sequence 6, Application US/08868786
 ; Patent No. 5998701

GENERAL INFORMATION:
 APPLICANT: Kawchuk, Lawrence M.
 APPLICANT: Armstrong, John
 APPLICANT: Lynch, Dermot
 APPLICANT: Knowles, Richard
 TITLE OF INVENTION: Potatoes Having Improved Quality
 TITLE OF INVENTION: Characteristics and Methods for Their Production
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 STREET: 5730 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: USA

ZIP: 80803
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/868,786
 FILING DATE: 04-JUN-1997
 CLASSIFICATION: 800

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/036,946
 FILING DATE: 10-FEB-1997

ATTORNEY/AGENT INFORMATION:
 NAME: Sullivan, Sally A
 REGISTRATION NUMBER: 32,064
 REFERENCE/DOCKET NUMBER: 8-97

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 974 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-868-786-6

Query Match 10.2%; Score 67; DB 2; Length 974;
 Best Local Similarity 21.1%; Pred. No. 31;
 Matches 30; Conservative 25; Mismatches 49; Indels 38; Gaps 5;

QY 14 LALLAAPAFQVPTVDLLANPPAGWISYGOQENYRHSPLTQIT 60
 Db 579 LCVVSGHVNAGVAIEHSEIVDEVFYKLMPEFQNKINGVPRRWLSGNCNELSEIT 638

QY 61 TENY-----QLOLVWARGMQPGKVQVTPPLIHGVMYLANPGDV 99

Db 639 TKWTGSDDWLVNTEKLAELRKFADNEELOSEWRKAKGNKMIIVSLIREKGYVSP-DA 697

QY 100 IQAIDAKTGDLIWEHRLPNI 121
 Db 698 MEDVQIKR--IHEKROLLNI 716

RESULT 11
 US-08-510-215A-2
 ; Sequence 2, Application US/08510215A
 ; Patent No. 5814506

GENERAL INFORMATION:
 APPLICANT: KONG, Huimin
 APPLICANT: PELLETIER, John J.
 APPLICANT: ALIOTTA, Jason M.
 TITLE OF INVENTION: OVER-EXPRESSION AND PURIFICATION OF
 TITLE OF INVENTION: A TRUNCATED THERMOSTABLE DNA POLYMERASE BY PROTEIN
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND
 STREET: 32 TOZER ROAD
 CITY: BEVERLY
 STATE: MASSACHUSETTS
 COUNTRY: US

ZIP: 01915
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/510,215A
 FILING DATE: 02-AUG-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: WILLIAMS, GREGORY D.
 REGISTRATION NUMBER: 30901
 REFERENCE/DOCKET NUMBER: NEB-113

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (508) 927-5054
 TELEFAX: (508) 927-1705

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 876 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-510-215A-2

Query Match 10.0%; Score 65.5; DB 2; Length 876;
 Best Local Similarity 23.0%; Pred. No. 40;
 Matches 38; Conservative 16; Mismatches 54; Indels 57; Gaps 7;

QY 14 LALLAAPA-----FAOVPTVDLLANPPAGWISYGOQENYRHSPLTQITTE 62
 Db 282 LERMAAPFAEGEKPLEMEFAIVDYITTEMLADKAA--LVEVMEENVHDAPIVGIALV 338

QY 63 N-----VGLOLVW-----ARGMQPGKVQVTPLIHGV-----M 91
 Db 339 NEHGRFEMPETALADSCFLAMLADETKKSMEDAKRAVVALKMGIELRGVAFDILLAA 398

QY 92 YLANP-----GDVQALDAKTGDLIWEH-----RROLPNITATL 124
 Db 399 YLLNPAQDAGDIAAAMKAKOYEAVRSDEAVYKGYKRSJPDQTL 443

RESULT 12
 US-09-310-293-2
 ; Sequence 2, Application US/09310293


```

US-08-852-091-2
; Sequence 2, Application US/08852091
; Patent No. 5883228
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 851 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-091-2

```

Query Match 9.9%; Score 65; DB 2; Length 851;

Best Local Similarity 22.9%; Pred. No. 44;
Matches 30; Conservative 16; Mismatches 43; Indels 42; Gaps 6;

```

QY 1 MKPISLMAAGALALAPAFQVPTVDELANPPAGW-----ISYQ 46
    | : | | | | : | | | | : | | | | : | | | | : |
DB 459 MNOISIMASVLAENLSPNLQND-----QFTSNPKAPWSLGPALSMQFSSYVGR 512
    | : | | | | : | | | | : | | | | : | | | | : |
QY 47 NOENYRHSPLTOITTEVQ-----LQLYA-----RGMOPKRV-----QVTPLIHDG 89
    | : | | | | : | | | | : | | | | : | | | | : |
DB 513 NSDQ-----LSMLRNKLFQGNCRTEPDLISWADFTKRESPKRLPFTWLDKLTLELVHDH 567
    | : | | | | : | | | | : | | | | : | | | | : |
QY 90 VMYLANPGDVI 100
    | : | | | | : | | | | : | | | | : | | | | : |
DB 568 LKDIMNDGRIM 578

```

Search completed: May 24, 2002, 10:02:56
Job time: 230 sec

10-12

10-13

10-14

10-15

10-16

10-17

10-18

10-19

10-20

10-21

10-22

10-23

10-24

10-25

10-26

10-27

10-28

10-29

10-30

10-31

10-32

10-33

10-34

10-35

10-36

10-37

10-38

10-39

10-40

10-41

10-42

10-43

10-44

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:02 ; Search time 56.04 Seconds
(without alignments)
214.332 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPTSLIMASAGALILALAP.....KTGDLIWEHRQLPNATLNL 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	28.9	742	2 A49340	alcohol dehydrogen
2	183	27.9	738	2 S14270	alcohol dehydrogen
3	178	27.1	742	2 JS0326	alcohol dehydrogen
4	165.5	25.2	708	2 AS2317	glucose dehydrogen
5	152.5	23.2	796	2 AG0523	glucose dehydrogen
6	150.5	22.9	796	1 JV0107	glucose dehydrogen
7	150.5	22.9	796	2 H90644	glucose dehydrogen
8	150.5	22.9	796	2 H85495	glucose dehydrogen
9	148	22.5	778	2 G98221	glucose dehydrogen
10	148	22.5	778	2 A13064	glucose dehydrogen
11	137	20.9	623	2 B83399	glucose dehydrogen
12	136.5	20.8	803	2 F83360	glucose dehydrogen
13	134.5	20.5	601	2 E95863	glucose dehydrogen
14	134	20.4	808	1 QPKEX	glucose dehydrogen
15	116	17.7	801	1 B09343	glucose dehydrogen
16	105.5	16.1	809	2 B98318	glucose dehydrogen
17	105.5	16.1	809	2 JQ0706	glucose dehydrogen
18	101.5	15.4	626	2 JQ0706	glucose dehydrogen
19	94.5	14.4	809	2 AS5547	glucose dehydrogen
20	93.5	14.2	639	2 JC4881	glucose dehydrogen
21	83.5	12.7	2802	2 F97686	glucose dehydrogen
22	83.5	12.7	2802	2 A12911	glucose dehydrogen
23	82	12.5	668	2 C75264	glucose dehydrogen
24	81	12.3	613	2 F69424	glucose dehydrogen
25	80.5	12.3	647	2 C83012	glucose dehydrogen
26	78.5	11.9	148	1 S03638	glucose dehydrogen
27	77	11.7	393	2 AD0350	glucose dehydrogen
28	77	11.7	733	2 S78376	glucose dehydrogen
29	76.5	11.6	400	2 AD0604	glucose dehydrogen

30	76	11.6	306	2 D75562	N-acetyl-gamma-glu
31	76	11.6	386	2 A82284	conserved hypotet
32	76	11.6	573	2 S68591	methanol dehydrog
33	76	11.6	580	2 AB0994	gamma-glutamyltran
34	76	11.6	772	2 H86016	hypothetical prote
35	75.5	11.5	224	2 AD2115	phosphoribosylform
36	75.5	11.5	407	2 C82804	conserved hypotet
37	75.5	11.5	684	2 T36771	probable integral
38	75.5	11.5	3972	2 S75251	hypothetical prote
39	74	11.3	380	2 C83171	conserved hypotet
40	74	11.3	969	2 A75634	McRb-related prote
41	73.5	11.2	584	2 AH2506	ser/thr protein ki
42	73	11.1	355	2 E95890	probable sugar upt
43	73	11.1	407	2 AG0821	probable lipoprote
44	73	11.1	407	2 G87515	secreted protein,
45	72	11.0	244	2 D75505	hypothetical prote

ALIGNMENTS

RESULT 1
A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence-revision 07-Apr-1994 #text-change 20-Jun-2000
C:Accession: A49340
J:Takekura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
R: Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:9517067; PIDN:BA40252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyridoxal-5-phosphate-quinone) (pyridoxal-5-phosphate-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 28.9% Score 190; DB 2; Length 742;
Best Local Similarity 38.8% Pred. No. 3.2e-11;
Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALILAPAFQVTPYDEL--ANPPAGEMISYGOENRHSPLTQTITENVGOLQ 68
DB 23 AALPYAAPARADGQNTGEAIIHADHPENWLSYGTYSQGRSPIDQINRSWVGLK 82
QY 69 LVMARGMOPGKVO-VTPLIHDSVMTLANPGDYIAIDAKTGDLWEHRQLP-NIA 122
DB 83 LAMVYTLDTNRGQENPLVVDGIMVATTWMSKEMALDADATGKLWQYDPKVGNTA 138

RESULT 2
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence-revision 26-Jul-1996 #text-change 20-Jun-2000
C:Accession: S14270
J:Takekura, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BA00528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyridoxal-5-phosphate-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MNR>

Query Match 27.9%; Score 193; DB 2; Length 738;

Best Local Similarity 33.3%; Pred. No. 1.6e-10;

Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALALAAPAFQVTPVDE-----LLANPPAGEWISYSGONGENRHSPLQTIT 61
 DB 17 TAGTICALISGATMASADGATGEGALIHADHPGNMWTGRTYSDDRTSPLDQINR 76
 QY 62 ENVGOLQVWARGMOPGVQV-TPLIHGVMTLANPGDVIOAIDAKTGLIWEHRQLP- 119
 DB 77 SNVGMKLAWYIDITDNRGQECTPLVDGVMATTNMAMKAVDAATGKLMSYDPRVPG 136
 QY 120 NTA 122
 DB 137 NTA 139

RESULT 3

JS0326

alcohol dehydrogenase (EC 1.1.-.-) 72k chain precursor - Acetobacter aceti

C:Species: Acetobacter aceti

C>Date: 31-Mar-1990 #sequence, revision 31-Mar-1990 #text, change 20-Jun-2000

C:Accession: JS0326

R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.

J. Bacteriol. 171, 3115-3122, 1989

A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase sub

A:Reference number: JS0326; MUID:89255070

A:Accession: JS0326

A:Molecule type: DNA

A:Residues: 1-742 <TNO>

A:Cross-references: GB:D90004; GB:M26951; NID:g216193; PIDN:BAA14058.1; PID:g216194

A:Experimental source: strain K6033

A>Note: amino terminal of mature protein is confirmed

C:Genetics:

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

F:1.35/Domain: signal sequence #status predicted <SIG>

F:36-742/Product: alcohol dehydrogenase 72k dehydrogenase chain #status predicted <ADC>

Query Match

Best Local Similarity 37.9%; Score 178; DB 2; Length 742;

Matches 44; Conservative 19; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALAAPAFQVTPVDELL-ANPPAGEWISYSGONGENRHSPLQTITENVGOLQ 68
 DB 23 AALPVAAYPARADGGGNGTGEALIHADHPENMLSTGRTYSDDRTSPLDQINRNVGDLK 82
 QY 69 LVWARGMOPGVQV-TPLIHGVMTLANPGDVIOAIDAKTGLIWEHRQLP-NIA 122
 DB 83 LIGVYTLIDNRGQECTPLVDGVMATTNMAMKAVDAATGKLMSYDPRVPGNIA 138

RESULT 4

S52317

quinomemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos

C:Species: Comamonas testos

C>Date: 08-May-1995 #sequence, revision 21-Jul-1995 #text, change 02-Jun-2000

C:Accession: S62366; S62373; S65908; S52317

R:Stoorvogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 235, 690-696, 1996

A:Title: Characterization of the gene encoding quinomemoprotein ethanol dehydrogenase c

A:Reference number: S62366; MUID:96184549

A:Accession: S62366

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-708 <STOL>

A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CAA57464.1; PID:g663196

A:Status: preliminary

A:Molecule type: protein

A:Residues: 32-54 <STO2>

R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 230, 899-905, 1995

A:Title: Quinomemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific

A:Reference number: S65908; MUID:95324580

A:Accession: S65908

A:Molecule type: protein

A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>

A:Experimental source: ATCC 15667

C:Genetics:

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; quinoprotein

F:1.31/Domain: signal sequence #status predicted <SIG>

F:32-708/Product: quinomemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match

Best Local Similarity 25.2%; Score 165.5; DB 2; Length 708;

Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PSLWASAGALALAAPAFQVTPVDE-----DELLANPP-AGEWISYSGONGEN 50
 DB 12 PGRMVLILACLG--SAAAFQGTGPAQAAAVQRYDGFIRANARPPDPPTGVDAE 69
 QY 51 YRHSPLQTITENVGOLQVWARGMOPGVQV-TPLIHGVMTLANPGDVIOAIDAKTGD 109
 DB 70 TRYSRLDQINRNVGDLKGLAMSYNLESTRGVATPVDGIMYVSAHSVVAIDITRIGN 129
 QY 110 LIWEHRQL 118
 DB 130 RIWTTDPQI 138

RESULT 5

AG0523

glucose dehydrogenase (imported) - Salmonella enterica subsp. enterica serovar Typh

C:Species: Salmonella enterica subsp. enterica serovar Typh

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence, revision 09-Nov-2001 #text, change 27-Nov-2001

C:Accession: AG0523

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr

, S.; Moule, S.; O'Gaora, P.

Mature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AG0523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:g16501455; GSPDB:GNO0176

C:Genetics:

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 23.2%; Score 152.5; DB 2; Length 796;

Best Local Similarity 33.3%; Pred. No. 2.3e-07;

Matches 43; Conservative 17; Mismatches 44; Indels 25; Gaps 5;

QY 6 LTVMS-----AGALALAAPAFQVTPVDELLANPPAGEWISYSGONGENRHSPLQT 58
 DB 135 LTVMGFNDPQELNGLTADATPA-ATSSSIAD-----DMPAYGRNQGGRVSPKQ 185
 QY 59 ITTENVGOLQVWARGMOPGVQV-TPLIHGVMTLANPGDVIOAIDAKTGD 109
 DB 186 ITADNVHQLKAWFRGDLKQPNDRPQELTNEVTPKIKVGTLYLCTAHQRTALDAASGR 245
 QY 110 LIWEHRQL 118

Db 246 EKMHPDL 254

RESULT 6

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Escherichia coli*

C:Species: *Escherichia coli*

C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000

C:Accession: D64735; J0107; A45997; S45201; I41228

R:Blatter, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: D64735

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-796 <BLAT>

A:Cross-references: GB:AE000122; GB:U00096; NID:g1786315; PIDN:AACT3235.1; PID:g1786316;

A:Experimental source: strain K-12, substrain M6165

R:Clifton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.

J. Bacteriol. 172, 6308-6315, 1990

A:Title: Cloning, mapping, and sequencing of the gene encoding *Escherichia coli* quinop

A:Reference number: J0107; MUID:91035240

A:Accession: J0107

A:Molecule type: DNA

A:Residues: 1-36, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>

A:Cross-references: GB:X51323; NID:941553; PIDN:CA35706.1; PID:941554

A:Experimental source: strain K12

R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.

J. Biol. Chem. 268, 12812-12817, 1993

A:Title: Topological analysis of quinoprotein glucose dehydrogenase in *Escherichia coli*

A:Reference number: A45997; MUID:93286127

A:Accession: A45997

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <YAM>

R:Fujita, N.

submitted to the EMBL Data Library, January 1994

A:Reference number: S45181

A:Accession: S45201

A:Molecule type: DNA

A:Residues: 1-796 <PUJ>

A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BA05580.1; PID:g473791

A:Experimental source: strain K-12 substrain W3110

R:Yamada, M.; Asoka, S.; Saiter, M.H.; Yamada, Y.

J. Bacteriol. 175, 568-571, 1993

A:Title: Characterization of the *gcd* gene from *Escherichia coli* K-12 W3110 and regulatio

A:Reference number: I41228; MUID:93123180

A:Accession: I41228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>

A:Cross-references: GB:D12651; NID:g216555; PIDN:BA02174.1; PID:g216556

C:Genetics:

A:Gene: *gcd*

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

Query Match 22.9%; Score 150.5; DB 1; Length 796;

Best Local Similarity 31.8%; Pred. No. 3,7e-07; Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;

Db 135 LTMAGFNDPQELINGTSLADATPAEA-ISPVDQ-----DMPAYGRNQGGRFSLKQ 185

OY 6 LTMAS-----AGALALLAPAFACVTPVDELANPPAGEWISYQNGENYRHSPLTQ 58

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H85495

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobleck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85495

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <STO>

A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AA654428.1; GSPDB:GN00145; UMGCP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: *gcd*

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

Query Match 22.9%; Score 150.5; DB 2; Length 796;

Best Local Similarity 31.8%; Pred. No. 3,7e-07; Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;

Db 135 LTMAGFNDPQELINGTSLADATPAEA-ISPVDQ-----DMPAYGRNQGGRFSLKQ 185

OY 6 LTMAS-----AGALALLAPAFACVTPVDELANPPAGEWISYQNGENYRHSPLTQ 58

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H85495

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobleck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85495

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <STO>

A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AA654428.1; GSPDB:GN00145; UMGCP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: *gcd*

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

A:Gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 22.9%; Score 150.5; DB 2; Length 796;
Best Local Similarity 31.8%; Pred. No. 3.7e-07;
Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;

QY 6 LMAS-----AGALALAAPAFQVPTVDELALNPAGGEMISGQNGENYRHSPLTQ 58
DB 135 LTMAGNDPQENGTISALATRAEA-ISPVADQ-----DWPAYGRNQESQRESPLKQ 185
QY 59 ITTENVGOLQVLA-----RGMPGKV--QVPLIHGVMYLANPGDVAIDAKTGD 109
DB 186 INADVNHNLEKAMVFTGDKQPNPDGETTEVTPIKGDTLYLCTAQRLEFALDAASGK 245
QY 110 LIMEHRROL 118
DB 246 EKMWHPDEL 254

RESULT 9
G98221
glucose dehydrogenase PA2290 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98221
R:Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: G98221
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-778 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK69297.1; PID:q15159131; GSPDB:GN00170
A:Gene: AGR_L1436
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 22.5%; Score 148; DB 2; Length 778;
Best Local Similarity 30.0%; Pred. No. 6.5e-07;
Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

QY 3 PLSLMA-SAGALALLAAPAF-----QVPTVDELALNP-----GEMI 42
DB 108 PTGLSTSPNAMPVLVSLASFVAGYSMAQDPHDQAGSLPQETASAPVYGGVPPGDWH 167
QY 43 SYGQNGENYRHSPLTQITTEVNGOLQVLAARGMPGKV-----QVPLIHGVM 91
DB 168 QYGTTPYGRYRSPLTQVNVNDSQLKEAMR--YQTGDVLPDDVGETTYQVPLKIGNTL 225
QY 92 YLANPGDVAIDAKTGDILMEHRROLPNI 121
DB 226 YICPPHMAIADAAATGKKMKYD---PNV 252

RESULT 10
A13064
glucose dehydrogenase gcd [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: A13064
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, S.; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; ster, E.W. The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193
A:Accession: A13064
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-778 <KUR>
A:Cross-references: GB:AE008689; PIDN:AA144935.1; PID:q17742589; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Gene: gcd
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 22.5%; Score 148; DB 2; Length 778;
Best Local Similarity 30.0%; Pred. No. 6.5e-07;
Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

QY 3 PLSLMA-SAGALALLAAPAF-----QVPTVDELALNP-----GEMI 42
DB 108 PTGLSTSPNAMPVLVSLASFVAGYSMAQDPHDQAGSLPQETASAPVYGGVPPGDWH 167
QY 43 SYGQNGENYRHSPLTQITTEVNGOLQVLAARGMPGKV-----QVPLIHGVM 91
DB 168 QYGTTPYGRYRSPLTQVNVNDSQLKEAMR--YQTGDVLPDDVGETTYQVPLKIGNTL 225
QY 92 YLANPGDVAIDAKTGDILMEHRROLPNI 121
DB 226 YICPPHMAIADAAATGKKMKYD---PNV 252

RESULT 11
B83399
quinoprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83399
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L. Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: B83399
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-623 <SNO>
A:Cross-references: GB:AE004624; GB:AE004091; NID:9947973; PIDN:AG05370.1; GSPDB:GN00187
A:Experimental source: strain PA01
C:Gene: exaA; PA1982

Query Match 20.9%; Score 137; DB 2; Length 623;
Best Local Similarity 31.5%; Pred. No. 6.4e-06;
Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;

QY 3 PLSLMA-SAGALAL--LAAPAFQVPTVDELALNP--PRAGEMISGQNGENYRHSPLT 57
DB 9 PAGLLRPLHCLAFVAVAGSAGALAKYVTMEDIANDDKTTGTDVLTQMGTHAQRMSPLK 68
QY 58 QITTEVNGOLQVLAARGN---QPGKYQVTPPLIHGVMYLANPGDVAIDAKTGDILME 113
DB 69 QVNADNVEKFLPAMVSFGDEKRGQ--ESQATVSDGVYIYVAYSRLFALDAKGRKLTW 127
QY 114 HRROLPN 120
DB 128 YNHRLPD 134

RESULT 12
F83360
glucose dehydrogenase PA2290 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83360
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: AB2950; MUID:20437337
 A:Accession: F83360
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-803 <STO>
 A:Cross-references: GB:AE004654; GB:AE004091; NID:g9948311; PIDN:AG05678.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: gcd; PA2290
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 20.8%; Score 136.5; DB 2; Length 803;
 Best Local Similarity 27.4%; Pred. No. 1e-05;
 Matches 43; Conservative 17; Mismatches 50; Indels 47; Gaps 5;

QY 6 LWM-----ASAGATLAAAPFAQVTPYTDLLANP----- 36
 DB 103 LMLPFRPRLADGPAPLGTALGVAVVLAGAAAGSG-FTNPQIVGRIDRDSGMTST 161
 QY 37 -----PAGEWISYQONQENRHSPLTQITENYQLOLTVARMOQPKV-----Q 81
 DB 162 AAPMPGDMQAYGRTEFGGRYSPLKQITPANTGOLEAMR--IRGDLPTADPLELINE 219
 QY 82 VPLIHGVWYLANPQGVYQIADTGLIWEHRQL 118
 DB 220 NTEPLKVMGMVACTASHKVLADPTGAEIWFEPQI 256

RESULT 13
 E95863
 alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - *Sinorhizobium*
 C:Species: *Sinorhizobium meliloti*
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
 C:Accession: E95863
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A>Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95863
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-601 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48573.1; PID:g15140045; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A86039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb20173
 A:Genome: plasmid
 C:Keywords: oxidoreductase

Query Match 20.5%; Score 134.5; DB 2; Length 601;
 Best Local Similarity 33.0%; Pred. No. 1.1e-05;
 Matches 38; Conservative 27; Mismatches 41; Indels 9; Gaps 7;
 QY 14 LALLAAPAFQVTPYTDL--LANPAGWISYQONQENRHSPLTQITTEWQLOLW 71
 DB 113 EHRQL 118
 DB 248 VYDPKL 253

DB 8 LAINSIGGAOAVANDELQKLIIDP-NQWAIQIGDIYANLRYSLDQINKDNQKLOVAM 66
 QY 72 --ARGMQPKVQVPLIHGVWYLANP-GDYIQAID-AKTGDLIWEHR-ROLPNI 121
 DB 67 TFSGVGLNGH-EGSLVYIGDLMYHTPPNIVYALDLSKQDIYWKKEPKQDPNV 120

RESULT 14
 QPKEX
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda
 C:Species: *Gluconobacter oxydans*
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Cleton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 Mol. Gen. Genet. 229, 206-212, 1991
 A>Title: A single amino acid substitution changes the substrate specificity of quinop
 A:Reference number: S17716; MUID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CLE>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL data library, February 1992
 A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, A, 214-808 <GOO>
 A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417
 C:Genetics:
 A:Gene: gdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran
 F:9-38/Domain: transmembrane #status predicted <TM1>
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:121-138/Domain: transmembrane #status predicted <TM5>
 F:191-93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:470/Active site: Asp #status predicted

Query Match 20.4%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 1.8e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;

QY 13 ALALLAPAFQVTPYTDL-----NPPAGWISYQONQENRHSPLTQIT 61
 DB 128 AVLALFASLFTDPHDISGLPTQIANASPADPDVWPASEWHAAGRTQAGRWSPFLNDINA 187
 QY 62 ENVGQLOLW-----ARGMQPKV--QVPLIHGVWYLANPQGVYQIADTGLI 112
 DB 188 TNSVNLKAMHIHTRKMMNSNDPEQTNTEPIEFNTTLTMCSLHQKLFVADGATGNVKW 247

QY 113 EHRQL 118
 DB 248 VYDPKL 253

RESULT 15
 S00943
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Acinetobacter calco*
 C:Species: *Acinetobacter calcoaceticus*
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S00943
 R:Cleton-Jansen, A.M.; Goosen, N.; Ogle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A>Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenas
 A:Reference number: S00943; MUID:88289368
 A:Accession: S00943
 A:Molecule type: DNA

A:Residues: 1-801 <CLE>
A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712
A:Experimental source: strain LMD 79.41
A:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane F
F:9-53/Domain: transmembrane #status predicted <TM1>
F:39-57/Domain: transmembrane #status predicted <TM2>
F:61-79/Domain: transmembrane #status predicted <TM3>
F:94-108/Domain: transmembrane #status predicted <TM4>
F:118-137/Domain: transmembrane #status predicted <TM5>
F:91_93/Binding site: ubiquinone (Arg, Asp) #status predicted
F:471/Active site: Asp #status predicted

Query Match	17.7%	Score 116;	DB 1;	Length 801;
Best Local Similarity	29.8%;	Pred. No. 0.0012;		
Matches	31;	Conservative	14;	Mismatches 39;
			Indels	20;
			Gaps	3;

[illegible]

Search completed: May 24, 2002, 10:06:03
Job time: 287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:47 ; Search time 27.86 Seconds

(Without alignments)
173.724 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPSTLWNASGALALALAP.....KTGDLWEHRRQLPNIATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	206	31.4	757 1	DHET_GLUOX
2	183	27.9	738 1	DHET_ACEPO
3	182	27.7	739 1	DHET_ACEBU
4	178	27.1	742 1	DHET_ACEBU
5	150.5	22.9	796 1	DHET_ECOLI
6	137	20.9	623 1	EXXA_PSEAE
7	134	20.4	808 1	DHG_GLUOX
8	116	17.7	801 1	DHG_ACTCA
9	112.5	17.1	631 1	DHMI_PARDE
10	111.5	17.0	626 1	DHMI_METOR
11	101.5	15.4	626 1	DHMI_METEX
12	94.5	14.4	809 1	QUIA_ACTCA
13	91	13.9	600 1	XOXF_PARDE
14	90	13.7	790 1	QUIA_XANCU
15	87.5	11.9	148 1	C008_LODMI
16	77	11.7	733 1	PSAB_ODOSI
17	76	11.6	233 1	TNFA_MARMO
18	76	11.6	571 1	DHMI_METME
19	73	11.1	776 1	PSAB_HETTR
20	70.5	10.7	540 1	PPBI_RAT
21	70	10.7	441 1	GUN2_THERFU
22	70	10.7	441 1	YBHJ_ECOLI
23	69.5	10.6	753 1	DACC_ECOLI
24	69	10.5	400 1	INVI_CAPAN
25	68.5	10.4	417 1	HVM2_SELMA
26	68	10.4	347 1	YA99_MYCN
27	68	10.4	661 1	PSAB_PROHO
28	68	10.4	1515 1	GLTB_AZOB
29	67.5	10.3	146 1	AZUP_ALCPA
30	67.5	10.3	261 1	TPIS_MYCLE
31	67.5	10.3	524 1	P60_LISIV
32	67.5	10.3	734 1	PSAB_CYACA
33	67	10.2	159 1	MP63_MYCTU

ALIGNMENTS

RESULT	ID	Query Length	STANDARD	PRT	757 AA
1	DHET_GLUOX				
AC	005542				
DT	01-NOV-1997 (rel. 35, Created)				
DT	01-NOV-1997 (rel. 35, Last sequence update)				
DT	16-OCT-2001 (rel. 40, Last annotation update)				
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit I).				
GN	ADHA.				
OS	Glucobacter oxydans (Glucobacter suboxydans).				
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;				
OC	Glucobacter.				
OX	NCBI_Taxid=442;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.				
RC	STRAIN-IFO 12528;				
RX	MEDLINE-97208225; PubMed-9055427;				
RA	Kondo K., Horiuchi S.;				
RT	Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Glucobacter suboxydans and their expression in Acetobacter pasteurianus.				
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).				
RL	Appl. Environ. Microbiol. 63:1131-1138(1997).				
CC	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.				
CC	-1- COFACTOR: POQ AND HEME.				
CC	-1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL: D86375; BAA19753.1; -				
DR	HSSP: Q944J7; 1F1G				
DR	InterPro: IPR001479; Bac_POQ				
DR	InterPro: IPR002372; Bac_POQ_repeat				
DR	InterPro: IPR000345; Cyt_heme_bind				
DR	Pfam: PF01011; Bacterial_POQ_6				
DR	PROSITE: PS00363; BACTERIAL_POQ_1; 1				
DR	PROSITE: PS00364; BACTERIAL_POQ_2; 1				
DR	PROSITE: PS00190; CYTOCHROME_C_1				
KW	Oxidoreductase; POQ; Heme; Ferriplasmic; Membrane; Signal.				
FT	SIGNAL	1	34		
FT	CHAIN	35	757		ALCOHOL DEHYDROGENASE [ACCEPTOR]
FT	MOD_RES	35	35		PYRROLIDONE CARBOXYLIC ACID.
FT	ACT_SITE	342	342		BASE (POTENTIAL).

34	67	10.2	331 1	DIV_ECOLI
35	67	10.2	485 1	BIND_STREN
36	67	10.2	500 1	YFJ1_YEAST
37	67	10.2	864 1	STRA2_PIG
38	67	10.2	974 1	PHS2_SOLRU
39	66.5	10.1	1550 1	GLTB_SYNY3
40	66	10.0	346 1	GPDA_XYLA
41	66	10.0	369 1	YDHI_ECOLI
42	66	10.0	527 1	YF00_MYCPN
43	65.5	10.0	224 1	PUR0_SYNY3
44	65.5	10.0	330 1	RLX3_STAV
45	65.5	10.0	353 1	OM51_HAEIN

P15286	escherichia
P3118	strongyloce
P43601	saccharomyc
O02799	sus scrofa
P53535	solanum tub
P55037	synechocyst
O9PCH7	xyliella fas
P7570	escherichia
P75287	mycoplasma
O55843	synechocyst
P14491	staphylococ
P43840	haemophilus

DR PROSITE; PS00364; BACTERIAL_POO_2; 1.

DR	PROSITE; PS00363;
DR	PROSITE; PS00364;

DR PROSITE; PS00190; CYTOCHROME_C; 1.

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE-93286127; PubMed-8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: PO4.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X51323; CA35706.1; -;
 DR EMBL; D12851; BA02174.1; -;
 DR EMBL; D26562; CAB20298.1; -;
 DR EMBL; AE000122; AAC73235.1; -;
 DR PIR; J01017; J01017;
 DR HSSP; P38539; 4AAH;
 DR EcGene; EG10369; gcd.
 DR InterPro; IPR001478; Bac_PQO.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO_7.
 DR PROSITE; PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 RN
 FT DOMAIN 1 10
 FT TRANSSEM 11 37 CYTOPLASMIC (PROBABLE).
 FT TRANSSEM 38 40 PROBABLE.
 FT DOMAIN 39 47 PERIPLASMIC (PROBABLE).
 FT TRANSSEM 41 58 PROBABLE.
 FT TRANSSEM 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSSEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSSEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSSEM 119 141 PERIPLASMIC (PROBABLE).
 FT DOMAIN 142 147 PROBABLE.
 FT ACT SITE 146 466 PERIPLASMIC (PROBABLE).
 FT ACT SITE 59 59 BASE (PROBABLE).
 FT CONFLICT 149 156 R -> L (IN REF. 1).
 FT CONFLICT 193 193 TLSADATP -> HLKRCRT (IN REF. 1 AND 2).
 FT CONFLICT 666 666 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;

Query Match 22.9%; Score 150.5; DB 1; Length 796;
 Best Local Similarity 31.8%; Pred. No. 3.1e-07;

Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;
 Oy 6 LIMS-----AGALLAAPAQVPTVDELLANPPAGEWISGQONENRHSPLTQ 58
 Db 135 LTMAGFNDPDRINLTSLADATPAEA-ISPVDQ-----DMPAYGRNGGRFSLKQ 185
 Oy 59 ITTENVGOLQVMA-----RGMQPKV--QVPELHDGVMYLANGDVYQAIADATGD 109
 Db 186 INADNVHLKAWVFRIGDVAKQPDPELTNEVPIKVGDTLYCTAHRFLDALDASGK 245
 Oy 110 LIMEHRQL 118
 Db 246 EKWYDEPL 254
 RESULT 6
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 ID EXAA_PSEAE
 AC 092417;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Madman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schober M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-20202376; PubMed-10736230;
 RA Keitel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: PO4 AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.


```

CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AJ009858; CAA0896.1; -
CC EMBL: AE004624; AAC05370.1; -
CC EMBL: AF068264; AAC79657.1; -
CC PDB: 1PLG; 30-AUG-00.
CC InterPro: IPR001479; Bac_POO.
CC InterPro: IPR002572; Bac_POO_repeat.
CC Pfam: PF01011; Bacterial_POO; 7.
CC PROSITE: PS00363; BACTERIAL_POO_1; FALSE_NEG.
CC PROSITE: PS00364; BACTERIAL_POO_2; FALSE_NEG.
CC Oxidoreductase; POO; Periplasmic; signal; Calcium; 3D-structure;
CC Complete proteome.
CC KW SIGNAL 1 34 QUINOPROTEIN ETHANOL DEHYDROGENASE.
CC FT CHAIN 35 623
CC DISULFID 139 140
CC FT SEQUENCE 623 AA; 68123 MW; 32D5DF20B291D6 CRC64;
CC SQ

```

CC	-1	SUBCELLULAR LOCATION: OUTSIDE THE PERIPLASMIC MEMBRANE.
CC	-1	MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE. WHILE THE
CC	P2	FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
CC		HERE.
CC	-1	SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC		or send an email to license@sib-sib.ch).
CC		-----
CC	EMBL, X62710;	CAA44594.1; ALT_SEQ.
DR	PIR, S17716;	OPKEX.
DR	InterPro: IPR001479;	Bac_POO.
DR	InterPro: IPR002372;	Bac_POO_repeat.
DR	Pfam: PF01011;	Bacterial_POO: 7.
DR	PROSITE: PS00363;	BACTERIAL_POO.1; 1.
DR	PROSITE: PS00364;	BACTERIAL_POO.2; 1.
KW	Oxidoreductase, POQ;	Transmembrane; periplasmic; Signal.
FT	SIGNAL	1 33 POTENTIAL.
FT	CHAIN	34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE].
FT	TRANSMEM	35 54 POTENTIAL.
FT	TRANSMEM	59 76 POTENTIAL.
FT	TRANSMEM	94 108 POTENTIAL.
FT	TRANSMEM	123 138 POTENTIAL.
FT	ACT_SITE	470 470 HASE (POTENTIAL).
FT	VARIANT	788 788 H -> N (IN P2 FORM).
FT	SEQUENCE	808 AA; 87567 MW; 0F4160DA78652445 CRC64;

```

Query Match          20.4%; Score 134; DB 1; Length 808;
Best Local Similarity 27.8%; Pred. No. 1.3e-05;
Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps

OY      13 ALALLAAPAFQAVTPVDELLA-----NPPAGEMISYGQONENYRHSPLTOITTT 61
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       128 AVLAIFASLFTDPHDISEGLTQTQANMSPADPDVDVAPESEMHAYGRTOAGDRWSPLQINA 187
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      62 ENVGQLQLTW-----ARGMGPKV-QVTPLIHGCVMIANGDYIQALIDAKTGSLIW 112
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       188 TVNSLVKVAWMIHTKDMANSNDPGEOTNEATPIEFNNLTLYMCSLHOKLFPAVDGATGNVKW 247
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      113 EHRROL 118
           :   :   :
Db       248 VYDPKL 253

RESULT    8
DHGA_ACT1CA ID DHGA_ACT1CA STANDARD: PRT: 801 AA.
AC P05465;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
   (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDHA.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
CC NCBL_TaxID=471;
OX
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMD 79.41;
RX MEDLINE=88289368; PubMed=3399393;
RA Cleton-Jansen A.-M., Goosen N., Dale G., van de Putte P.;
de "Nucleotide sequence of the gene coding for quinoprotein glucose
RT dehydrogenase from Acinetobacter calcoaceticus." ;
RL Nucleic Acids Res. 16:6228-6228(1988).

```

CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERS
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07235; CAA30222.1; -.
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 FT ACT_SITE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
 SQ SEQUENCE

Query Match 17.7%; Score 116; DB 1; Length 801;
 Best Local Similarity 29.8%; Pred. No. 0.0008;
 Matches 31; Conservative 14; Mismatches 39; Indels 20; Gaps 3;

QY 20 PAFQAQVPTVDLPPAGEMWISYGNQENYRHSPLQTITENYQLOLVAMRGKQOPK 79
 DB 152 PETAQAVPGVAE-----SDMPAYGRGTQAGVRYSPKLOINDQNVADLKAAMT--LRTGD 202
 QY 80 V-----QVTPLIHGVVYLANPGVYQAIQDAKTGDLW 112
 DB 203 LKTDNDGSETTNOVTPIKIGNNMFICTAHQOLIAIDPATCKEKW 246

RESULT 9
 DHML_PARDE STANDARD; PRT; 631 AA.
 ID DHML_PARDE
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Paracoccus.
 CC NCB1_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RX MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.,
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RT structural gene from Paracoccus denitrificans.";
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 CC TOTAL CELL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17339; AAA88366.1; -.
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631
 FT DISULFID 135 136
 FT DISULFID 418 447
 FT ACT_SITE 335 335
 FT ACT_SITE 631 AA; 69799 MW; 0934DC93FEC5730B CRC64;
 SQ SEQUENCE

Query Match 17.1%; Score 112.5; DB 1; Length 631;
 Best Local Similarity 32.6%; Pred. No. 0.0014;
 Matches 42; Conservative 25; Mismatches 47; Indels 15; Gaps 9;

QY 4 TSLMASAGALL--AAPFAQVPTVDL--LANPAGEMWISYGNQENYRHSPLQIT 60
 DB 12 SSIAMAVMGLAVLTAPATA-----NDQVELANDPA--NMYMGRDINAQNTSEMDIN 65
 QY 61 TENYQLOLVW--ANGMDGKQVPLPIHGVVYLANP-GDVYQAIQDA-KTGDLWEHR- 115
 DB 66 KENVYQQLRPMSEFSYGVVGHG-EGTPLYGVDRMFITPFPYTTALDINERKILMOKR 124
 QY 116 ROLPIATL 124
 DB 125 KQNPATATV 133

RESULT 10
 DHML_METOR STANDARD; PRT; 626 AA.
 ID DHML_METOR
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Methylobacterium group; Methylobacterium.
 CC NCB1_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
 RX STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
 RC MEDLINE=89008094; PubMed=2459109;
 RA Machlin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 RT Methylobacterium organophilum XX methanol dehydrogenase structural
 RT gene.";
 RL J. Bacteriol. 170:4739-4747(1988).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

Q01A ACICA STANDARD: PRT: 809 AA.

AC 059086:

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Quinate/shikimate dehydrogenase [Pyridoxalpyridine-quinone]

DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).

GN Q01A.

OS Acinetobacter calcoaceticus.

OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;

OC Acinetobacter.

OX NCBI_TaxID=471;

RN [1]

RP SEQUENCE FROM N.A. AND CHARACTERIZATION.

RC STRAIN-BD413 / ADP1.

RA MEDLINE=95059536; PubMed=8002591;

RA Elsemore D.A., Ornston L.N.;

RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus

RT contains quina, the structural gene for quinate-shikimate

RT dehydrogenase";

RL J. Bacteriol. 176:7659-7666(1994).

RN [2]

RP SEQUENCE OF 1-20 FROM N.A.

RC STRAIN-BD413 / ADP1;

RX MEDLINE=96011389; PubMed=7592351;

RA Elsemore D.A., Ornston L.N.;

RT "Unusual ancestry of dehydrogenases associated with quinate catabolism

RT in Acinetobacter calcoaceticus";

RL J. Bacteriol. 177:5971-5978(1995).

CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.

CC -1- CATALYTIC ACTIVITY: quinate + pyridoxalpyridine-quinone = 5-

CC dehydroquinolate + reduced pyridoxalpyridine-quinone.

CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRIDOXALPYRIDINE-QUINONE = 3-

CC DEHYDROSHIKIMATE + REDUCED PYRIDOXALPYRIDINE-QUINONE.

CC -1- COFACTOR: PQQ.

CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY, FIRST STEP. THIS PATHWAY

CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO

CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE

CC PATHWAY.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).

CC -1- INDUCTION: BY PROTOCATECHUATE.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L05770; AAC37161.1; -

DR InterPro: IPR001479; Bac_PQQ.

DR InterPro: IPR002372; Bac_PQQ_repeat.

DR Pfam: PF01011; Bacterial_PQQ; 7.

DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.

DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.

KW Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.

FT TRANSMEM 14 34 POTENTIAL.

FT TRANSMEM 41 61 POTENTIAL.

FT TRANSMEM 68 88 POTENTIAL.

FT TRANSMEM 90 110 POTENTIAL.

FT TRANSMEM 127 147 POTENTIAL.

SO SEQUENCE 809 AA; 88196 MW; 71F67CBEA62BFCB CRC64;

Query Match 14.4%; Score 94.5; DB 1; Length 809;

Best Local Similarity 29.2%; Pred. No. 0.11;

Matches 35; Conservative 13; Mismatches 49; Indels 23; Gaps 4;

QY 3 PTLNWSAGALLAA-PAFAOVPTVDLLANPPAGWISYGCNGEYRHSPLTQIT 61

DB 148 PHEWKAAGEELPLVPDPACKOV-----NMHYGNDAGSRFVALDQINR 193

QY 62 ENVGLOLVW-----ARGNQPG-KVOVPLIHGVYLANPGDVIAIDAKTGDLWE 113

DB 194 NNVSRLKEAWRPRTGDTTGTGNGAEDQMTPLQNGKVFCTPPHNIFADIDSGKOLWK 253

RESULT 13

XOXF_PARDE

ID XOXF_PARDE STANDARD: PRT: 600 AA.

AC P29968;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).

GN XOXF.

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Paracoccus.

OX NCBI_TaxID=266;

RN [1]

RP SEQUENCE FROM N.A.

RA Harms N.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 380-600 FROM N.A.

RC STRAIN-PD 1235;

RX MEDLINE=92041583; PubMed=1657873;

RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,

RA Stouthamer A.H.;

RT "Isolation, sequencing, and mutagenesis of the gene encoding

RT cytochrome c551 of Paracoccus denitrificans and characterization of

RT the mutant strain.";

RL J. Bacteriol. 173:6971-6979(1991).

CC -1- COFACTOR: PQQ (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U34346; AAC44555.1; -

DR EMBL: M75583; AAA2574.1; -

DR PIR: A41378; A41378.

DR HSSP: P38539; 4AAH.

DR InterPro: IPR002372; Bac_PQQ_repeat.

DR Pfam: PF01011; Bacterial_PQQ; 7.

KW Oxidoreductase; PQQ; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.

FT ACT_SITE 318 318 BASE (POTENTIAL).

SO SEQUENCE 600 AA; 65159 MW; DCA9966F1BC5A3CE CRC64;

Query Match 13.9%; Score 91; DB 1; Length 600;

Best Local Similarity 28.2%; Pred. No. 0.17;

Matches 35; Conservative 22; Mismatches 45; Indels 22; Gaps 7;

QY 10 SAGALLAALAAFAOVPTVDLLANP-----PAGWISYGCNGEYRHSPLTQITTE 62

DB 6 NGACIALMSGTAA-----LANDRAGRDQAPQWMIQMGDYANTRYSTLDQINKD 56

QY 63 NVGQLQVW--ARGNQPGKVOVPLIHGVYLANP-GDVIQAIIDAK-TGDLWEHR-RQ 117

DB 57 NVKDLRAVMTFSTGLRGH-EGSPLVIGDVYVHTPPNRFALDLNDGKILMRYPQO 115

QY 118 LPNI 121

DB 116 DPNV 119

RESULT 14
 QUITA_XANCU STANDARD; PRT: 790 AA.
 ID QUITA_XANCU STANDARD; PRT: 790 AA.
 AC 09XD78:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable quinate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25).
 GN QUITA.
 OS Xanthomonas campestris (pv. juglandis).
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 OX NCBI_TaxID=44291;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C5:
 RA MEDLINE=20063481; PubMed=10594704;
 RA Lee Y.-A., Lo Y.-C., Yu P.-P.,
 RT "A gene involved in quinate metabolism is specific to one DNA homology
 RT group of Xanthomonas campestris."
 RL J. Appl. Microbiol. 87:649-658(1999).
 CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
 CC dehydroquinone + reduced pyrroloquinoline-quinone.
 CC -1- COFACTOR: POQ (BY SIMILARITY).
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETODIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF109471; AAD38453.1; -
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_6.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 CC Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 77 94 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 SQ SEQUENCE 790 AA; 82896 MW; B75F29B52A4FE6F CRC64;

Query Match 13.7%; Score 90; DB 1; Length 790;
 Best Local Similarity 27.9%; Pred. No. 0.29;
 Matches 36; Conservative 18; Mismatches 53; Indels 22; Gaps 6;

QY 9 ASAGALLA-LAFAAQTPTVDE---LLANPAG---EWISYQONENYRHSPLTOI 59
 DB 110 ATYATLAAVVGIGMFPVHPVAGNMGKTAVPQSGVQONMSAYGNTDGSFFALDOI 169
 QY 60 TTEWVGOLQVMAKMGPKGV-----QVTPLIHDSVMTLANPGDVIQADKTGD 109
 DB 170 NRSN-GRPAGSGPPTTPEGLANSDNGAGMDQLTPLOVGEKVFICPNNLIALDASTGK 228
 QY 110 LIWEHROL 118
 DB 229 QLV--RREI 235

RESULT 15
 CU08_LOCM1 STANDARD; PRT: 148 AA.
 ID CU08_LOCM1 STANDARD; PRT: 148 AA.
 AC P11734:
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cuticle protein 8 (LM-8) (LM-ACP 8).
 DE Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Phryganea; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=90073593; PubMed=2590176;
 RA Klarskov K., Hoelrup P., Andersen S.O., Roepstorff P.;
 RT "Plasma-desorption mass spectrometry as an aid in protein sequence
 RT determination. Application of the method on a cuticular protein from
 RT the migratory locust (Locusta migratoria)."
 RL Biochem. J. 262:923-930(1989).
 RN [2]
 RP SEQUENCE OF 1-56.
 RA MEDLINE=86108304; PubMed=3943519;
 RA Hoelrup P., Andersen S.O., Roepstorff P.;
 RT "Isolation, characterization, and N-terminal sequence studies of
 RT cuticular proteins from the migratory locust, Locusta migratoria."
 RL Eur. J. Biochem. 154:153-159(1986).
 CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
 CC CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
 CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-[AV]) REPEATS FOUND THROUGHOUT THE
 CC PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
 CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
 CC -1- SIMILARITY: CONTAINS 1 CUTICLE CONSENSUS DOMAIN.
 CC
 CC PIR: B24802; B24802.
 CC PIR: S05638; S05638.
 CC InterPro: IPR000618; Insect_cuticle.
 CC Pfam: PF00379; Insect_cuticle.
 CC PROSITE: PR00947; CUTICLE.
 CC PROSITE: PS00233; CUTICLE.
 CC structural protein; Cuticle; Repeat.
 FT REPEAT 16 19
 FT REPEAT 22 22
 FT REPEAT 28 31
 FT REPEAT 37 40
 FT REPEAT 44 47
 FT REPEAT 5.5
 SQ SEQUENCE 148 AA; 15224 MW; C6EADC27C593ACE6 CRC64;

Query Match 11.9%; Score 78.5; DB 1; Length 148;
 Best Local Similarity 23.1%; Pred. No. 0.61;
 Matches 24; Conservative 17; Mismatches 20; Indels 43; Gaps 3;

QY 8 WASAGALLA-LAFAAQTPTVDELLANPAGEWISYQONENYRHSPLTOITENVGOL 67
 DB 27 YAAVAKAVVAPAVAVAPAKAVVAP-----VAYPKYFENG----- 66
 QY 68 QLVMAKMGPKGVQVTPLIHDSVMTLANPGDVIQADKTGDI 111
 DB 67 -----VHD-----AHGDIKQSGEARDGDV 87

Search completed: May 24, 2002, 10:15:48
 Job time: 766 sec

Fri May 24 11:27:14 2002

us-08-934-506a-5_copy_1_125.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:28 ; Search time 93.93 Seconds

(without alignments)
230.218 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125
Perfect score: 657
Sequence: 1 MKPTSLIMASAGALIALIAP.....KTGDLIWEHRRLPNTATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254.5	38.7	608	2	Q93RE9
2	190	28.9	742	2	Q53362
3	178.5	27.2	691	2	Q9AF95
4	174	26.5	623	2	Q9AGN3
5	168.5	25.6	698	2	Q9KH03
6	167	25.4	786	2	P95466
7	165.5	25.2	708	2	Q46444
8	156	23.7	777	16	Q9F9U2
9	155	23.6	785	16	Q98KF6
10	146.5	22.3	777	16	Q92R33
11	136.5	20.8	803	16	Q91115
12	134.5	20.5	601	2	Q9EYW8
13	134.5	20.5	601	16	Q92WY9
14	121	18.4	790	2	Q9X255
15	118.5	18.0	695	2	Q934G0
16	112	17.0	601	2	P71509

17	110.5	16.8	629	2	Q9AQ48
18	103	15.7	633	2	Q24759
19	93.5	14.2	639	2	P77931
20	93.5	14.2	644	2	Q52551
21	89	13.5	499	2	Q51842
22	83.5	12.7	2832	2	Q93N05
23	83	12.6	599	2	Q9L935
24	82.5	12.6	485	17	Q96X7
25	82	12.5	668	16	Q9R8H3
26	81	12.3	613	17	Q28873
27	80.5	12.3	647	16	Q9HUB1
28	77	11.7	846	2	Q86092
29	76	11.6	306	16	Q9R772
30	76	11.6	386	16	Q9K7W8
31	76	11.6	573	2	Q59540
32	75.5	11.5	407	16	Q9PG38
33	75.5	11.5	684	2	Q9S282
34	75.5	11.5	3972	16	P73139
35	75	11.5	680	5	Q9N967
36	74	11.3	380	16	Q9HXJ7
37	74	11.3	525	16	Q931E9
38	74	11.3	969	16	Q9RZ15
39	73.5	11.2	183	16	Q9CKL1
40	73.5	11.2	1229	5	Q9VW05
41	73	11.1	355	16	Q926F4
42	73	11.1	407	16	Q9A6E2
43	72	11.0	244	16	Q9RWM1
44	72	11.0	392	2	Q9F5L2
45	72	11.0	506	11	Q9DAQ4

ALIGNMENTS

RESULT 1
ID Q93RE9 PRELIMINARY; PRT: 608 AA.
AC Q93RE9; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudomonas aeruginosa
OC Bacteria; Pseudomonadaceae.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shida T., Saito Y.;
RT "Alcohol dehydrogenase."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1;
SQ SEQUENCE 608 AA; 65101 MW; 0ACEC97AE11BA570 CRC64;

Query Match 38.7%; Score 254.5; DB 2; Length 608;
Best Local Similarity 41.5%; Pred. No. 5.2e-17;
Matches 54; Conservative 18; Mismatches 39; Indels 19; Gaps 1;
QY 10 SAGALIALIAPFAO-----VPTDELLANPAGEMISYGNQEN 50
DB 16 STALLASUSGPAFOHDANAAPSKAGOSALENFPVADDLAKGNPANWPLRBNYOG 75
QY 51 YHNSPLQITTTNVQQLVMARGMQPGVQVYTPLLHDSVMTLANPGDVIAQIDAKTGD 110
DB 76 WYSPLDQINKNVGLQVWSRTMEPGSNEGAAIYNGVIFLGNNTNDVIAQIDKGTSL 135
QY 111 IWEHRRLPNT 120
DB 136 IWEYRRKLPS 145

RESULT 2

Q53362 PRELIMINARY; PRT; 742 AA.

AC Q53362; Q44159;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)

DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.

OS Acetobacter pasteurianus (Acetobacter turbidans).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

OC Acetobacter.

OX NCBI_TaxID=438;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NC11380;

RA MEDLINE=94042848; PubMed=822628;

RA Takemura H., Kondo K., Horinouchi S., Beppu T.;

RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteurianus";

RL J. Bacteriol. 175:6857-6866(1993).

DR EMBL; D13893; BAA40252.1; -.

DR HSSP; Q924J7; 1FLG.

DR InterPro: IPR001479; Bac_PQO.

DR InterPro: IPR002372; Bac_PQO_repeat.

RT Pfam: PF01011; Bacterial_PQO; 6.

DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.

DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.

SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B813EC98 CRC64;

Query Match 28.9%; Score 190; DB 2; Length 742;

Best Local Similarity 38.8%; Pred. No. 1.8e-10;

Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGAALLAPAFAGVPTVDL--ANPAGEMISYGNQENYRHSPLQTTEWVGOL 68

DB 23 AALPAPVAPARADGGNGEALIHADHPENWLSIGRTYSQKRSPLDQIRSNVGD 82

QY 69 LVWARGMPGKVO-VTPIIHGVMYLANPGDIVIOAIDAKTGLIWEHRRLP-NIA 122

DB 83 LAMVYLTDTNRGEATPLVVDGIMVATWMSKMEALDAAATGKLWQYDPKVGNTA 138

RESULT 3

Q9AF95 PRELIMINARY; PRT; 691 AA.

AC Q9AF95;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE 1-BUTANOL DEHYDROGENASE BDH.

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=86174;

RN [1]

RP SEQUENCE FROM N.A.

RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora";

RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF355796; AAK27220.2; -.

SQ SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 27.2%; Score 178.5; DB 2; Length 691;

Best Local Similarity 37.8%; Pred. No. 2.3e-09;

Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

QY 39 GEMISYGNQENYRHSPLQTTEWVGOLVWARGMPGK-VQVTPLIHGVWYLANPG 97

DB 41 GEMRTHGDDAGTRISPLAQITPDNAKELGLWYSTDESSRGVEATPIVVDGVMIVTAFW 100

QY 98 DVIOAIDAKTGLIWEHRRLP 119

DB 101 SVVHALDVRSGRKRLWTYDPEVP 122

RESULT 4

Q9AGW3 PRELIMINARY; PRT; 623 AA.

AC Q9AGW3;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=86174;

RN [1]

RP SEQUENCE FROM N.A.

RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora";

RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF326086; AAK15506.1; -.

DR HSSP; Q924J7; 1FLG.

DR InterPro: IPR002372; Bac_PQO_repeat.

DR Pfam: PF01011; Bacterial_PQO; 4.

KW Signal.

FT SIGNAL.

SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 26.5%; Score 174; DB 2; Length 623;

Best Local Similarity 38.0%; Pred. No. 5.6e-09;

Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLMASAGALLAPAFAGVPTVDL--ANPAGEMISYGNQENYRHSPLQT 59

DB 7 KPFALRAIVATAALASLPA-AAVTDVTDWIDANDHKTGDDVLTGLGAQSHSPLKAI 65

QY 60 TTEWVGOLQVMA-----RCMQPGKVOVTPPIIHGVMYLANPGDIVIOAIDAKTGLI 111

DB 66 NTDVNVANLVPAWSFSFGKRGQGE--AGV--LVHDGVYIATASYSRIFAIDARSGRL 120

QY 112 WEHRRLP 120

DB 121 WEYNAKLPD 129

RESULT 5

Q9KH03 PRELIMINARY; PRT; 698 AA.

AC Q9KH03;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-DEC-2000 (TREMBLrel. 15, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE TETRAHYDROFURFURYL ALCOHOL DEHYDROGENASE.

OS Alcaligenes eutrophus (Ralstonia eutropha).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI_TaxID=510;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21125557; PubMed=11222593;

RA Zarrt G., Schrader T., Andreessen J.R.;

RT "Catalytic and Molecular Properties of the Quinohemoprotein Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha Strain Bo.";

RT J. Bacteriol. 183:1954-1960(2001).

DR EMBL; AF277373; AAF86335.1; -.

DR HSSP; Q924J7; 1FLG.

FT	TRANSMEM	701	721	POTENTIAL.
FT	DOMAIN	722	786	PERIPLASMIC (POTENTIAL).

CC COLLMANOUAS, C. S. 1999, *Chem. Rev.* 99:1732-1734 (2001).
 CC ACTA CRYSTALLOGR. D 57:1732-1734 (2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF
 CC THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.

CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X61880; CAA57464.1; .
 DR HSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C.
 DR Signal: PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT TYPE I.
 FT BINDING 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACCAB3 CRC64;

Query Match 25.2%; Score 165.5; DB 2; Length 708;
 Best Local Similarity 31.8%; Pred. No. 4.6e-08;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

OY 3 PSLIMASAGALALAPAFAPVPTD-----DELLAMP-AGENISYQONEN 50
 DB 12 PGRWVLLAACLG--SAAAFQTGPAQAQAAAVQVDCDFRANAARTPTTGIVYAE 69
 OY 51 YRHSPLQITTEYNGQLVWARGMPQK-VQVPLIHGVMYLANGVYQAIIDAKTGD 109
 DB 70 TRYSLLDINANVKDGLAWSYNLESTRGVEATPVVDGIMYASMSVYHAIDRTGN 129
 OY 110 LIMEHRRQL 118
 DB 130 RIMTYDPOI 138

RESULT 8
 OY9F902 PRELIMINARY; PRT; 695 AA.
 AC OY9F902;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 GN ALCOHOL DEHYDROGENASE.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2.;
 RL Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; .
 DR HSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 23.7%; Score 156; DB 2; Length 695;
 Best Local Similarity 33.6%; Pred. No. 4e-07;

Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;
 OY 15 ALLAPAFAPVPTDELL--ANPAGEMISYQONENYRSPITQITTEWVGOLOIYWA 72
 DB 16 ALLVA-AGAQAQAKVDEAIRASEQDSEMLSHGRYARQRRSPKQIDAGVNGKGLAWY 74
 OY 73 RMOPK-VQVPLIHGVMYLANGVYQAIIDAKTGD LIMEHRRQL 118
 DB 75 LDLENKRGLEATPVSDGLVYASLSWSRYMAVDLRSGKRLMQDPQV 121

RESULT 9
 OY98KF6 PRELIMINARY; PRT; 785 AA.
 AC OY98KF6;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DE GLUCOSE DEHYDROGENASE.
 GN MLI1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=MAFE303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Matsumoto M., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti".
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002997; BAB48838.1; .
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 KW Complete proteome.
 SQ SEQUENCE 785 AA; 84083 MW; 1DC224B7A1C25470 CRC64;

Query Match 23.6%; Score 155; DB 16; Length 785;
 Best Local Similarity 31.7%; Pred. No. 5.8e-07;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;
 OY 8 WASAG-----ALLAPAFAPVPTD-----ELLANPAGE 40
 DB 111 WARRGLAGPDRAPLILAVLSAVAGYSMTADPKGIDGALDNDKVIIPKANLGNDVPAE 170
 OY 41 WISYQONENYRSPITQITTEYNGQLVWARGMPQK-----QVPLIHG 89
 DB 171 WHYGRQDFGQRYSPDLQIPPDVANLQPAWT--YRTGVKGGPDIGETTYQVPLKIGD 228
 OY 90 VMYLANGVYQAIIDAKTGD LIMEH 114
 DB 229 TLVICTHFNRAIAYDAATGKEKMY 253

RESULT 10
 OY92RB3 PRELIMINARY; PRT; 777 AA.
 AC OY92RB3;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYRROLOQUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

[illegible][illegible][illegible][illegible][illegible][illegible]

Fri May 24 11:27:15 2002

RT fixing endosymbiont Sinorhizobium meliloti.
NL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
KM Oxidoreductase; plasmid: Hypothetical protein; Complete proteome.
SQ SEQUENCE 601 AA; 65759 MW; D/3424FEED13ADBE CRC64.

Query Match
Best Local Similarity 20.58; Score 134.5; DB 16; Length 601;
Matches 38; Conservative 27; Mismatches 41; Indels 9; Gaps 7;

OY 14 LALLAPAPAVTEVDEL--LAMPAGEMISYGNQENYRHSPLTITTEWVGOLQVW 71
DB 8 LAINISGGAQVAFANDELOFLDDP-NOMAIQNGDIANIRYSKLDINKNVGKLOVAM 66
OY 72 --ARGMQPKVQVYPLIHDSVMTLAMP-GDVIOAID-ATGDLIWEH-NOLPNI 121
DB 67 TFESTVLNGH-EGSPLYVDLMTVHTPPVYVALDSNDQIVWKYEPEKDDPNV 120
RESULT 14
ID 09X255
AC 09X255
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DE 01-JUN-2001 (TREMBLrel. 12, Last sequence update)
GN GLUCOSE DEHYDROGENASE.
OS Pantoea citrea.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=53336;
RA MEDLINE=9958;
RX STRAIN=1056;
RT "Ujol C.J. 303331; PubMed=10376838;
RL "gdb, a g, Kado C.I.
RM Pantoea citrea, encoding a second glutoprotein glucose dehydrogenase in
DR EMBL, AF050503.45:1217-1226(1999).
DR Interf. Pro. IPR001A023735.1
DR Pfam: IPR002379; Bac. POC
DR PROSITE: PS00364; Bacterial POC repeat.
SQ SEQUENCE 790 AA; 85681 MW; 88BB7625E0A5F089 CRC64.

Query Match
Best Local Similarity 18.48; Score 121; DB 2; Length 790;
Matches 31; Conservative 13; Mismatches 39; Indels 12; Gaps 3;

OY 27 PVTDELLANPRAGEMISYGNQENYRHSPLTITTEWVGOLQVW 71
DB 152 PATQPVSAK--LMTAYGGTQAGDRTSSLSQINESVKNLQVAMTYOSGVKRPDDIGE 208
OY 80 V--VTPLIHGQVYLANPGDVIOAIDAKTKGDLI 112
DB 209 ITNESTLKVNLVTCRAHOILVAIDATNGKEK 243
RESULT 15
ID 0934G0
AC 0934G0
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
GN LUPANINE HYDROXYLASE PRECURSOR.
OS Pseudomonas sp. DH2001
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

us-08-934-506a-5-copy_1_125.rpt

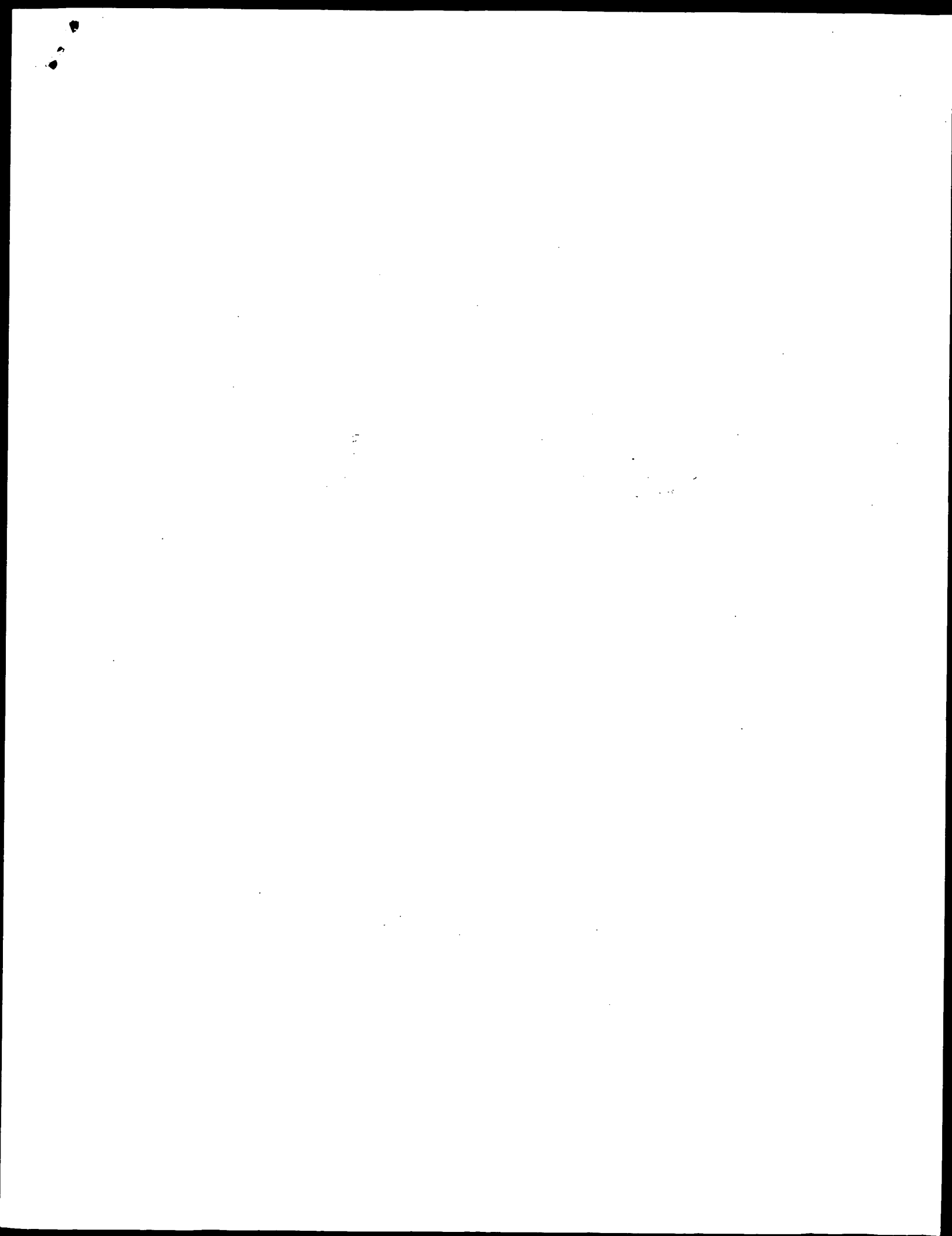
OC Pseudomonas.
OX NCBI_TaxID=163360;
RN 11
RP SEQUENCE FROM N.A.
RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
RT "Clon
RM hydroxylase, sequencing and analysis of the gene for lupanine
DR Submit 4 (JUN-2001) to the EMBL/GenBank/DBJ databases.
KW EMBL: A318095; CAC07410.1;
FT SIGNAL
FT CHAIN
SQ SEQUENCE 27 1 26
ID 96
AC 96
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
GN LUPANINE HYDROXYLASE.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

Query Match
Best Local Similarity 18.08; Score 118.5; DB 2; Length 695;
Matches 28; Conservative 8; Mismatches 41; Indels 1; Gaps 1;
OY 38 AGEWISYGNQENYRHSPLTITTEWVGOLQVWARGMQPKVQV--TPLIHDSVMTLAMP 96
DB 36 SGWMSLLGSGNEOHYFSAHLSPLTITTEWVGOLQVWARGMQPKVQV--TPLIHDSVMTLAMP 96
OY 97 GDVIOAIDAKTKGDLI 114
DB 96 PKIYANDLKTIKRNLMY 113

Search completed: May 24, 2002, 10:17:30
Job time: 743 sec

Fri May 24 11:27:15 2002

us-08-934-506a-5_copy_1_125.rspt



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:00 ; Search time 117.28 Seconds

(without alignments)
121.227 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128

Perfect score: 673
Sequence: 1 MKPSTLWMSAGALLALAAP.....DLWEHRRQLPNATLNSFG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	100.0	579	19	AAW37873
2	603	89.6	579	19	AAW37874
3	552	82.0	579	19	AAW37876
4	550.5	81.8	578	19	AAW37875
5	189.5	28.2	754	21	AAW35987
6	183	27.2	738	12	AAW35987
7	183	27.2	738	12	AAW35987
8	177	26.3	742	11	AAW35987
9	155	23.0	443	22	AAW35987
10	145.5	21.6	740	20	AAW35987
11	129	19.2	948	22	ABG25909

12	129	19.2	1510	22	ABG21573
13	76	11.3	824	18	AAW09614
14	73.5	10.9	1229	22	ABB71736
15	72.5	10.8	343	19	AAW35001
16	72.5	10.8	462	17	AAW05395
17	72	10.7	282	22	AAW1927
18	71	10.5	526	22	AAW9895
19	70	10.4	1016	22	AAW34349
20	69.5	10.3	275	21	AAW42781
21	69.5	10.3	1029	22	ABG26194
22	69.5	10.3	1029	22	ABG26568
23	69.5	10.3	1419	22	ABW35208
24	69	10.3	711	22	ABB69743
25	69	10.3	957	22	ABB62401
26	68.5	10.2	353	22	AAW4447
27	68.5	10.2	384	22	AAW85255
28	68.5	10.2	551	22	AAW85251
29	68.5	10.2	551	22	AAW85252
30	68.5	10.2	843	20	AAV34485
31	68.5	10.2	849	20	AAV34360
32	68	10.1	159	19	AAW63034
33	68	10.1	159	21	AAW14325
34	68	10.1	294	21	AAW35811
35	68	10.1	294	21	AAW15001
36	68	10.1	424	22	AAW59716
37	67.5	10.0	171	22	AAU32330
38	67.5	10.0	275	21	AAW16656
39	67.5	10.0	735	22	ABG24896
40	67.5	10.0	974	19	AAW40079
41	67.5	10.0	1041	22	AAW21594
42	67.5	10.0	1224	18	AAW26735
43	67	10.0	668	22	AAW66214
44	67	10.0	673	22	ABW59674
45	67	10.0	1407	22	ABG28517

ALIGNMENTS

RESULT 1	AAW37873	standard; Protein; 579 AA.
AC	AAW37873;	
AC	AAW37873;	
DE	10-AUG-1998 (first entry)	
XX	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulononic acid; L-ascorbic; inhibition.	
OS	Glucanobacter oxydans.	
XX	Key	Location/Qualifiers
FT	Peptide	1..23
FT	Protein	/note="signal peptide"
FT		24..579
FT		/note="mature protein"
XX	EP832974-A2.	
XX	PD	01-APR-1998.
XX	11-SEP-1997;	97EP-0115801.
PR	19-SEP-1996;	96EP-0115001.
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS
 XX Claim 1; Pages 35-37; 59pp; English.
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 100.0%; Score 673; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8.3e-66;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALALLAFAFAQVTPVDELLANPPAGEWISYGNONEVRRSPILQIT 60
 Db 1 mkptslimwasagalalalaaapaafaqvtpvdeellanppagewisynqenyrnsplqlt 60
 QY 61 TENVGQLQLVWARGMQPKVQVTPPLIHGVMYLANPGDVYQAIDAKTGDLIWEHRQLPN 120
 Db 61 tenvgqlqlvwargmqpkvqvtpplihgvmylanpgdviqaidaktgdliwehrrqlpn 120
 QY 121 IATLNSFG 128
 Db 121 iatlinsfg 128

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein; 579 AA.
 AC AAM37874;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 FH Key
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 EN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS
 XX Claim 1; Pages 38-40; 59pp; English.
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 89.6%; Score 603; DB 19; Length 579;
 Best Local Similarity 91.4%; Pred. No. 4.7e-58;
 Matches 117; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALALLAFAFAQVTPVDELLANPPAGEWISYGNONEVRRSPILQIT 60
 Db 1 mkptslimwasagalalalaaapaafaqvtpvdeellanppagewisynqenyrnsplqlt 60
 QY 61 TENVGQLQLVWARGMQPKVQVTPPLIHGVMYLANPGDVYQAIDAKTGDLIWEHRQLPN 120
 Db 61 tenvgqlqlvwargmqpkvqvtpplihgvmylanpgdviqaidaktgdliwehrrqlpn 120
 QY 121 IATLNSFG 128
 Db 121 iatlinsfg 128

RESULT 3
 AAM37876
 ID AAM37876 standard; Protein; 579 AA.
 AC AAM37876;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 FH Key
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 EN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

PT production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 XX Claim 1; Fig 8; 96pp; English.
 XX
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from *Glucobacter oxydans*. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially *Glucobacter* with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 XX Sequence 754 AA:

Query Match 28.2%; Score 189.5; DB 21; Length 754;
 Best Local Similarity 39.0%; Pred. No. 3.9e-12;
 Matches 46; Conservative 17; Mismatches 46; Indels 9; Gaps 3;

OY 14 LLLAAPAFAYTPV-----TDELLANPPAGEWISYQONENYRHSPLQITTEENYGO 66
 Db 17 lgcAaalfatcspalaeatgtaltndqbpqdmgyrlysegryplditkdnasn 76
 OY 67 LQLVWARGMQPGKQYV-TPLIHDGVMYLANPGDVIAQIDAKTGDLIWEHRRLP-NIA 122
 Db 77 lklawhyldltmrgegrplvdygmyatltwsmkkaidaatgkllwsydkvpqnia 134

RESULT 6
 AAR13993
 ID AAR13993 standard; Protein: 738 AA.

XX AC AAR13993;
 XX DT 09-DEC-1991 (first entry)
 XX DE A.altoacetigenes membrane-bound ADH 72kd sub-unit.
 XX KW alcohol dehydrogenase complex; carboxylic acid production.
 XX OS Acetobacter altoacetigenes.
 XX EP448969-A.
 XX PN 02-OCT-1991.
 XX PD 26-FEB-1991; 91EP-0102793.
 XX PF 26-FEB-1991; 90JP-0073440.
 XX PR 26-FEB-1990; 90JP-0073440.
 XX PR 26-FEB-1990; 90JP-0042391.
 XX PA (NAKA-) NAKANO VINEGAR KK.
 XX PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;
 XX PI Kawamura Y;
 XX DR WPI: 1991-289462/40.
 XX DR N-PSDB; AAQ13580.
 XX
 PT Gene for membrane-bound alcohol dehydrogenase complex - optd.
 PT from Acetobacter altoacetigenes, used for prodn. of enzyme for
 PT converting alcohol to acid
 XX
 PS Disclosure; Fig 3; 36pp; English.
 XX
 CC Total DNA was prepared from A.altoacetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation

CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. altoacetigenes (see AAQ13582-Q13584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 CC
 XX Sequence 738 AA:

Query Match 27.2%; Score 183; DB 12; Length 738;
 Best Local Similarity 33.3%; Pred. No. 2e-11;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGALALAAPAFAYTPVDE-----LANPPAGEWISYQONENYRHSPLQIT 61
 Db 17 tagticaallsgyatmasaddggatgaalnhaddhpgnmwlygrtysdqgrpldqlnr 76
 OY 62 ENVGQLVWARGMQPGKQYV-TPLIHDGVMYLANPGDVIAQIDAKTGDLIWEHRRLP- 119
 Db 77 snvgnlklawylldltmrgegrplvdygmyatltwsmmkkaidaatgkllwsydkvpqnia 136
 OY 120 NIA 122
 Db 137 nla 139

RESULT 7
 AAR20192
 ID AAR20192 standard; Protein: 738 AA.

XX AC AAR20192;
 XX DT 16-APR-1992 (first entry)
 XX DE ADH complex protein (mol.wt. 72.000).
 XX KW Alcohol dehydrogenase; acetic acid; fermentation.
 XX OS Acetobacter altoacetigenes NH-24.
 XX PN JF0326998-A.
 XX PD 27-NOV-1991.
 XX PF 26-MAR-1990; 90JP-0073440.
 XX PR 26-FEB-1990; 90JP-0042301.
 XX PR 26-MAR-1990; 90JP-0073440.
 XX PA (NAKA-) NAKANO SUTEN KK.
 XX PI WPI: 1992-019325/03.
 XX PI N-PSDB; AAQ20383.
 XX
 PT Alcohol dehydrogenase complex structural gene - used in plasmid
 PT and enhancing efficiency of acetic acid fermentation for
 PT transformed acetic acid bacteria
 XX
 PS Disclosure; Fig 3(1-3); 21pp; Japanese.
 XX
 CC Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex
 CC can be easily extracted from the bacteria and purified and it can be
 CC used for the determination of an alcohol.
 CC See also AAQ20383-84, and -86-88.
 XX
 XX Sequence 738 AA:

Db 168 ltwagfndpqlngltlsadatpaea-ispvadq-----dwpaygrngqgrfispkq 218
 QY 59 ITTENYGOLOLWVA-----RGMQSGKY--QVTPLIHGVWYLANPGVIOAIDAKTGD 109
 Db 219 lnadrvhnlkeavvrtgtgkqpnopgcltnevpikvgdltlylctahqrlfaldaasqk 278
 QY 110 LIWEHRRQLPNIATLNSF 127
 Db 279 ekwhyd---pelktnesf 293

RESULT 10

AAW95019
 ID AAW95019 standard; Protein; 740 AA.

AC AAW95019;

DT 21-MAY-1999 (first entry)

DE Sorbitol dehydrogenase (SLDH) protein sequence.

KM Sorbitol dehydrogenase; SLDH; open reading frame; ORF2; L-sorbose;
 KW D-sorbitol; Vitamin C; enzyme.

OS Gluconobacter suboxydans.

EH Key Location/Qualifiers

FT Peptide 1..24
 /note= "signal sequence"

FT Protein 25..740
 /note= "mature protein"

PN EEP97984-A2.

PD 24-FEB-1999.

PF 13-AUG-1998; 98EP-0115231.

PR 21-AUG-1997; 97EP-0114432.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1999-134646/12.

DR N-PSDB; AAX21501.

PT New D-sorbitol dehydrogenase gene and recombinant protein - useful
 for production of L-sorbose, an intermediate in vitamin C production

PS Claim 1; Fig 3A-D; 39pp; English.

CC This represents a sorbitol dehydrogenase (SLDH) protein. The DNA
 encoding the SLDH enzyme also encodes an open reading frame (ORF2)

CC product upstream of the SLDH open reading frame, needed for SLDH activity
 in vivo. Host cells transformed by a vector comprising the SLDH DNA

CC sequence are used for the recombinant expression of the sorbitol
 dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for

CC producing L-sorbose from D-sorbitol. L-sorbose is an important
 intermediate in vitamin C production.

CC Sequence 740 AA;

Query Match 21.6%; Score 145.5; DB 20; Length 740;
 Best Local Similarity 34.0%; Pred. No. 2.8e-07;
 Matches 34; Conservative 25; Mismatches 30; Indels 11; Gaps 5;

QY 23 AOVTPVTDLILANPPAGWISYGOENCRHSPLQITTEYNGOLQ--VMAKGM--QPG 78
 Db 86 sqvpmappgsanpargdwvaygriddhqrtyrpslspenasklkvalvnyhtgyprrp 145
 QY 79 KV-----QVTPV-LHGVWYLANPGDVIOAIDAKTGDILW 112

Db 146 qvnxwaetpikvgdjltycsamndlik-ldpatgqdlw 184

RESULT 11

ABG25909
 ID ABG25909 standard; Protein; 948 AA.

AC ABG25909;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25900.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS90096.

PT New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity

PS Claim 20; SEQ ID No 56268; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes.

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
 disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pdb/published_pot_sequences.

CC Sequence 948 AA;

Query Match 19.2%; Score 129; DB 22; Length 948;
 Best Local Similarity 29.9%; Pred. No. 2.6e-05;
 Matches 43; Conservative 17; Mismatches 30; Indels 54; Gaps 8;

QY 6 LWMAS-----AGALLAAPAFQVTPVTDLILANPPAGWISYGOENCRHSPLQ 58

OY 71 WARGMOPGKQVTPILHD-----GVATLANP---GDVIO-AIDAKT----- 107
 Db 733 -----eygpeticprhrdefdpklptgekeevpgkpinpetgdvrvpvdsvtkypv 787
 OY 108 -GDLINEHRRLP 119
 Db 788 kgdsive-keelp 799

RESULT 14

ABE71736
 ID ABE71736 standard; Protein; 1229 AA.

AC ABE71736;

DE 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 42000.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15839.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

PS Disclosure; SEQ ID NO 42000; 21pp + Sequence Listing; English.

PS The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB16176-AB130511) and the encoded proteins.

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1229 AA;

RESULT 15
 ID AAW35001 standard; Protein; 343 AA.
 AC AAW35001;
 DE 21-MAY-1998 (first entry)
 DE Bacillus thermoloeovorans endoglucanase.
 KW Endoglucanase; cellulase; carboxymethylcellulose; cellulose;
 KW biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification;
 KW thermostable enzyme; thermophilic; glycosidase.
 OS Bacillus thermoloeovorans (Clone 68GCl).
 PN WO9744361-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; 97MO-US08793.
 PR 22-MAY-1996; 96US-0651572.
 PA (RECO-) RECOMBINANT BIOCATALYSIS INC.
 PI Lam DE, Mathur EJ;
 DR WPI; 1998-018435/02.
 DR N-PSDB; AAT94209.
 PT Endoglucanase(s), preferably form archaeal bacterium, AEP11a -
 PT useful to degrade carboxymethylcellulose and hydrolyse of
 PT beta-1,4-glycosidic bonds in cellulose
 PS Claim 1; Fig 1Q; 164pp; English.
 CC This protein comprises an endoglucanase of Bacillus thermoloeovorans
 CC (Clone 68GCl) that is capable of degrading carboxymethylcellulose
 CC and of hydrolysing the beta-1,4-glycosidic bonds in cellulose. It
 CC has homology to an endoglucanase of archaeobacterium AEP11a (see
 CC AAW34985). It can be produced from native cells or from recombinant
 CC host cells, especially prokaryotic host cells transformed with a
 CC plasmid or virus-derived vector including the endoglucanase DNA
 CC (see AAT94209). 24 Endoglucanases (see AAW34986-W35008) are claimed.
 CC They can be used to degrade cellulose for the conversion of plant
 CC biomass into fuels and chemicals, for use in detergents, textiles,
 CC animal feed, waste treatment, and in the fruit juice and brewing
 CC industries for the clarification and extraction of juices.
 XX Sequence 343 AA;

Query Match 10.8%; Score 72.5; DB 19; Length 343;
 Best Local Similarity 35.3%; Pred. No. 12;
 Matches 18; Conservative 6; Mismatches 26; Indels 1; Gaps 1;

OY 28 VTDELANPPAGEWISYGNONGENRHSPLTQITTEWVGLQVWARGMOPG 78
 Db 106 vlenasamplgegwekgvtltkwestfaldatttnehgkyrywag-kdpg 155

Search completed: May 24, 2002, 10:05:01
 Job time: 240 sec

OY 8 WMSAGLALLAPAFQV--TPVTDELANP--PAGEWISYGNONGEN--YRHSPLTQITTE 62
 Db 958 yvskpavalsaaepalavplvshstgspclgygvgygssnggvallgpalckltsa 1017
 OY 63 NV-----GOLQVWARGMOPG 78
 Db 1018 payalggkiasstayaigag 1037

Fri May 24 11:27:15 2002

us-08-934-506a-5_copy_1_128.rag


```

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frischauf, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about
IDENTIFICATION METHOD: 72,000
ORIGINAL SOURCE:
ORGANISM: Acetobacter altoacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshiaki;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Tayama, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horinouchi, Sueharu and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
TITLE: polyoxogenes
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

```

Query Match 27.2%; Score 183; DB 1; Length 738;
Best Local Similarity 33.3%; Pred. No. 5,7e-13;
Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGLALLAFAFAQVPTVDE-----LLANPPAGWISYQONENRHSPLPQIT 61
DB 17 TAGTICALLISGIVATMASADGAGATGATGAIHADHPGNNMTYGRYTSIDRYSPLDQINR 76

```

OY 62 ENVGLOLWARGNQPGKVY-TPLIHGVWYLANPGDVIOAIDAKTGDLIWEHRRLP- 119
DB 77 SNVGNLKIATWLDITNRGEGTPLVIDGVYATTTNMSKMAVADATGKILMSTDPKPVG 136
OY 120 NIA 122
DB 137 NIA 139

```

```

RESULT 3
US-09-296-284-25
Sequence 25, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Glucanobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Glucanobacter suboxydans
US-09-296-284-25

```

Query Match 26.8%; Score 180.5; DB 4; Length 720;
Best Local Similarity 40.4%; Pred. No. 1.1e-12;
Matches 40; Conservative 18; Mismatches 36; Indels 5; Gaps 3;

```

OY 26 TPVTDELLANPPAGWISYQONENRHSPLQITTEENQGLQLVARGNQPGKVY-TP 84
DB 5 TATN--ADQHPGDMWSYGRYTSSEORYSPLDQITKDNASNLKLAHWYDLDTNRGEGT 61
OY 85 LIHGVWYLANPGDVIOAIDAKTGDLIWEHRRLP-NIA 122
DB 62 LIYDGVYATTTNMSKMAIDATGKILMSTDPKPVGNIA 100

```

```

RESULT 4
US-09-136-251-2
Sequence 2, Application US/09136251A
Patent No. 6127156
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: OIYMA, Setsuko
APPLICANT: SHINOH, Masako
APPLICANT: TOMIYAMA, No. 61271561bun1
TITLE OF INVENTION: D-Sorbitol Dehydrogenase GENE
FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
CURRENT APPLICATION NUMBER: US/09/136,251A
CURRENT FILING DATE: 1998-08-19
EARLIER APPLICATION NUMBER: EP 97114432.4
EARLIER FILING DATE: 1997-08-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 739
TYPE: PRT
ORGANISM: Glucanobacter suboxydans
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(24)
US-09-136-251-2

```


RESULT 8
 PCT-US93-02172-11
 Sequence 11: Application PC/TUS9302172
 GENERAL INFORMATION:
 APPLICANT: La Jolla Cancer Research Foundation
 TITLE OF INVENTION: RECOMBINANT CALF INTESTINAL ALKALINE
 PHOSPHATASE
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: La Jolla Cancer Research Foundation
 STREET: 10901 North Torrey Pines Road
 CITY: La Jolla
 STATE: California
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version 1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/02172
 FILING DATE: 19930310
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/849,219
 FILING DATE: 10-MAR-1992
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 455-6480
 TELEFAX: (619) 455-0181

2IP: 10111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796.792
FILING DATE: 06-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/011,364
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hone, William J.
REGISTRATION NUMBER: 26,739
REFERENCE/DOCKET NUMBER: 07763/03301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-765-5070
TELEFAX: 212-258-2291
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
S-08-796-792-2


```

; Patent No. 6117662
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Traini, Christopher M.
; APPLICANT: Mathie, Thomas B.
; TITLE OF INVENTION: Respiratory nitrate reductase alpha
; FILE REFERENCE: GM10000
; CURRENT APPLICATION NUMBER: US/09/310,293
; CURRENT FILING DATE: 1999-05-12
; EARLIER APPLICATION NUMBER: 60/086,579
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1229
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-310-293-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65.5; DB 3; Length 1229;
Matches 33; Conservative 22; Mismatches 62; Indels 25; Gaps 8;

```

```

QY 2 KPTSLM--ASAGALALAAPAFQ-VTPVDELLANPPAGWIS---YGCNQENYRH- 53
DB 739 KPEITWREETGKLDIVVSLDFRMTATPLVSDIVL--PAATWYKHKHLSSTDMHPYVHP 796
QY 54 -----SPLTQITTE-----NNGQLQVWARKMOPG---KVQVPLIHGQVYLANPGDVI 100
DB 797 FNPALDPLMESRSDMDIYKTLAKAFSEMAKDYLPGETFKDVYVTPPLSHDTKQEIETSPYGVV 856
QY 101 QALDAKTGDLMEHRROLPNIA 122
DB 857 K--DMSKGEIAPVPGRTMPNFA 876

```

```

RESULT 13
US-09-579-376-2
; Sequence 2, Application US/09579376
; Patent No. 6316237
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Traini, Christopher M.
; APPLICANT: Mathie, Thomas B.
; TITLE OF INVENTION: Respiratory nitrate reductase alpha
; FILE REFERENCE: GM10000
; CURRENT APPLICATION NUMBER: US/09/579,376
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/310,293
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/086,579
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1229
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-579-376-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65.5; DB 4; Length 1229;
Matches 33; Conservative 22; Mismatches 62; Indels 25; Gaps 8;
QY 2 KPTSLM--ASAGALALAAPAFQ-VTPVDELLANPPAGWIS---YGCNQENYRH- 53
DB 739 KPEITWREETGKLDIVVSLDFRMTATPLVSDIVL--PAATWYKHKHLSSTDMHPYVHP 796

```

```

QY 54 -----SPLTQITTE-----NNGQLQVWARKMOPG---KVQVPLIHGQVYLANPGDVI 100
DB 797 FNPALDPLMESRSDMDIYKTLAKAFSEMAKDYLPGETFKDVYVTPPLSHDTKQEIETSPYGVV 856
QY 101 QALDAKTGDLMEHRROLPNIA 122
DB 857 K--DMSKGEIAPVPGRTMPNFA 876

```

```

RESULT 14
US-08-369-796-2
; Sequence 2, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 851 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-369-796-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65; DB 1; Length 851;
Matches 30; Conservative 16; Mismatches 43; Indels 42; Gaps 6;
QY 1 MKPTSLMASAGALALALAAPAFQVTPVDELLANPPAGW-----ISYQ 46
DB 459 NNGQLSIAWAVLWFLNLSPLNQO-----QFFSNPKAPWSLGLPALSMQFSSYVGRGL 512
QY 47 NOENRHSPLTQITTEYNGQ-----LQLYVA-----RGMOPGKV-----QVPLIHG 89
DB 513 NSDQ-----LSMLNKLFGQNCRTEDPLSWADFTKRESPPGKLPFTWIDKILELVHDH 567
QY 90 VMYLANPGDVI 100
DB 568 LKDLMDGRIM 578

```

```

RESULT 15

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:03 ; Search time 56.04 Seconds
(without alignments)
219.476 Million cell updates/sec

Title: US-08-934-506A-5_COPY_1_128
Perfect score: 673
Sequence: 1 MKPTSLIMASGALALLAAP.....DLIWEHRRLPNTATINSE 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	190	28.2	742	2 A49340	alcohol dehydrogen
2	183	27.2	738	2 S14270	alcohol dehydrogen
3	178	26.4	742	2 JS0326	alcohol dehydrogen
4	165.5	24.6	708	2 S52317	quinohemoprotein e
5	155	23.0	796	1 JV0107	glucose dehydrogen
6	155	23.0	796	2 H90644	glucose dehydrogen
7	155	23.0	796	2 H85495	glucose dehydrogen
8	154	22.9	796	2 AG0523	glucose dehydrogen
9	148	22.0	778	2 G98221	glucose dehydrogen
10	148	22.0	778	2 A13064	glucose dehydrogen
11	137.5	20.4	803	2 F83360	quinoprotein alcoh
12	137	20.4	623	2 B83399	alcohol dehydrogen
13	134.5	20.0	601	2 E95863	glucose dehydrogen
14	134	19.9	808	1 OPRKX	glucose dehydrogen
15	121.5	18.1	801	1 S00943	glucose dehydrogen
16	106.5	15.8	809	1 B98314	probable quinate d
17	106.5	15.8	809	2 A12968	hypothetical prote
18	101.5	15.1	626	2 JQ0706	alcohol dehydrogen
19	94.5	14.0	809	2 A55547	quinate-shikimate
20	93.5	13.9	639	2 JC4881	polyvinyl-alcohol
21	83.5	12.4	2802	2 F97686	cyclic beta-(1-2)
22	83.5	12.4	2831	2 A12911	beta (1-->2) gluc
23	82	12.2	668	2 C75264	probable serine/th
24	81.5	12.1	393	2 AD0350	probable lipoprote
25	81	12.0	613	2 F69424	conserved hypotet
26	80.5	12.0	647	2 C83012	probable chemotaxi
27	78.5	11.7	509	2 S05638	cuticle protein 8
28	78.5	11.7	509	2 C70624	probable two compo
29	77	11.4	733	2 S78376	photosystem I P700

30	76.5	11.4	224	2 AD2115	phosphoribosylform
31	76.5	11.4	400	2 AD0604	D-alanyl-D-alanine
32	76	11.3	306	2 D75562	N-acetyl-gamma-glu
33	76	11.3	386	2 A82284	conserved hypotet
34	76	11.3	573	2 S68591	methanol dehydroge
35	76	11.3	580	2 AB0394	gamma-glutamyltran
36	76	11.3	772	2 H86016	hypothetical prote
37	75.5	11.2	407	2 C82804	conserved hypotet
38	75.5	11.2	684	2 T36771	probable integral
39	75.5	11.2	3972	2 S75251	hypothetical prote
40	74	11.0	380	2 C83171	conserved hypotet
41	74	11.0	969	2 A75634	McR-related hypotet
42	73.5	10.9	584	2 AH2506	ser/thr protein k1
43	73	10.8	355	2 E95990	probable sugar upt
44	73	10.8	392	2 AG0821	probable lipoprote
45	73	10.8	407	2 G87515	secreted protein,

ALIGNMENTS

RESULT 1
A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:95117067; PIDN:BA40252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 28.2%; Score 190; DB 2; Length 742;
Best Local Similarity 38.8%; Pred. No. 4.9e-11;
Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALLAAPARQVPTVDLL--ANPPAGEMISYGCQNGEYRHSPLQITENVGOLD 68
DB 23 AALPFAAVPARADGGGNGEALIHADHPENWLSYGRYSQRYSPLDQINRSVGDILK 82
QY 69 LVMARGMOPGKYV-VPLIHGVMYLANPGDYIOADTGDILWEHRRLP-NTA 122
DB 83 LAMVYTLIDNRGGQENTPLVVDGIMVATMTMSKMEALDATGKLMQYDKVGNIA 138

RESULT 2
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter pasteurianus
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S14270
R:Takemura, T.; Fukaya, M.; Takemura, H.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BA400528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAY>

Query Match 27.2%; Score 193; DB 2; Length 738;
Best Local Similarity 33.3%; Pred. No. 2.5e-10;
Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALALAAPAFQVTPVTE-----LLANPAGEMISYSGONGENRHSPLTQIT 61
DB 17 TAGTICALISGVATFMSADSDGATGEAIHADHPGMMWRTGRTYSQRTSPUDQINR 76
QY 62 ENVGOLQVWARGMOPGKQV-TPLIHGVMYLANPGDVIQAIKATGDLIWEHRQLP 119
DB 77 SNVGNKILAMYLDDLTNRQEGTPLVIGVMYATTNMSKMAVDAATGKLMSYDPRVPG 136
QY 120 NIA 122
DB 137 NIA 139

RESULT 3
J50326
alcohol dehydrogenase (EC 1.1.-.-) 72k chain precursor - Acetobacter aceti
C:Species: Acetobacter aceti
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Jun-2000
C:Accession: J50326
R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.
J. Bacteriol. 171, 3115-3122, 1989
A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase subunit
A:Reference number: J50326; MUID:89255070
A:Accession: J50326
A:Molecule type: DNA
A:Residues: 1-742 <IMO>
A:Cross-References: GB:D90004; GB:M26951; NID:9216193; PIDN:BAAI4058.1; PID:9216194
A:Experimental source: strain K6033
A:Note: amino terminal of mature protein is confirmed
C:Genetics:
A:Gene: adh1
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-742/Product: alcohol dehydrogenase 72k dehydrogenase chain #status predicted <ADC>

Query Match 26.4%; Score 178; DB 2; Length 742;
Best Local Similarity 37.9%; Pred. No. 7.9e-10;
Matches 44; Conservative 19; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALAAPAFQVTPVTEDEL--ANPAGEMISYSGONGENRHSPLTQITENYGOLG 68
DB 23 AALPFAAVPARADGQGTGEAIHDDHDENNLSTYRTSEORFPLDQINSNVGDLK 82
QY 69 LVNARGMOPGKQV-TPLIHGVMYLANPGDVIQAIKATGDLIWEHRQLP-NIA 122
DB 83 LGGYTLDTNRQEGTPLVVDGIMYATTNMSKMAIDAAATGKLMLQYDPRVGNIA 138

RESULT 4
S52317
quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
C:Species: Comamonas testosteroni
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
C:Accession: S52317; S52373; S55908; S52317
R:Stoorvogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 235, 690-698, 1996
A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase C
A:Reference number: S52366; MUID:96184549
A:Accession: S52366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STOI>
A:Cross-References: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
A:Accession: S52373

A:Status: preliminary
A:Molecule type: protein
A:Residues: 32-54 <STOI>
R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 230, 899-905, 1995
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific
A:Reference number: S55908; MUID:95324380
A:Accession: S55908
A:Molecule type: protein
A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>
A:Experimental source: ATCC 15667
C:Genetics:
A:Gene: qheh
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; quinoprotein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 24.6%; Score 165.5; DB 2; Length 708;
Best Local Similarity 31.8%; Pred. No. 1.4e-08;
Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PTLNWSAGALALAAPAFQVTPVTE-----DELLANP-AGEMISYSGONGEN 50
DB 12 PGRWVWMLNACLG--SAAFAQGTGPAQAQAAVQVDSFIRNARTDWPFGVDYAE 69
QY 51 YRSPLTQITTEVNGOLQVWARGMOPGK-VQVTPLIHGVWYLANPGDVIQAIKATGCD 109
DB 70 TRYSRLDQINAAVWVNDGLAMSYNLESTRGVATEPVVDGIMVYSKMSVVAIDTRGDN 129
QY 110 LIWEHRQL 118
DB 130 RIWTPDQI 138

RESULT 5
J50107
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli
C:Species: Escherichia coli
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64733; J50107; A45997; S45201; I41228
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A:Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <BLAT>
A:Cross-References: GB:A6000122; GB:U00096; NID:91786315; PIDN:AACT3235.1; PID:917863
A:Experimental source: strain K-12, substrain MG1655
R:Cleton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino
A:Reference number: J50107; MUID:91035240
A:Accession: J50107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRGHT', 157-192, 'K', 194-796 <CLE>
A:Cross-References: GB:X51323; NID:941553; PIDN:CAA35706.1; PID:941554
R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia co
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181

A:Gene: STY0191
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match
Best Local Similarity 32.6%; Score 154; DB 2; Length 796;
Matches 45; Conservative 19; Mismatches 46; Indels 28; Gaps 6;

OY 6 LMAS-----AGALALAAPAFACVPTVDLLANPPAGEWISYQNGENYRHSPLTQ 58
DB 135 LSWAGFNDPQEIINGTLRADATPA-ATSSIADE-----DMPAYGRNQGQRYSPKQ 185
OY 59 ITTENYGOLOLWMA-----RGMPGKY--QVTPLIHDGVYKLANPGDVIQDAIDAKGD 109
DB 186 ITRDNYHQLKAMVFERGTGLKQPNDEITTEVTPIKVGDTLICTAHQRFALDAASGK 245
OY 110 LIWEHRRLPNIATLNSF 127
DB 246 EKMHPD---POLKTDSSF 260

RESULT 9
G98221
glucose dehydrogenase PA2290 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98221
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: G98221
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-778 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89297.1; PID:g15159131; GSPDB:GN00170
C:Genetics:
A:Gene: AGR.L.1436
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match
Best Local Similarity 22.0%; Score 148; DB 2; Length 778;
Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

OY 3 PTSLMA-SAGALALAAPAF-----QVTPVDLLANPPA-----GEMI 42
DB 108 PTGLSYSPNAMPPLVLSVASFVAVGYSMAODPHDQAGSLPQELISAPVYGEVPGDDMH 167
OY 43 SYGNGENYRHSPLTQITTEENVGLOLVWARGMPGKY-----QVTPLIHDGVY 91
DB 168 QYGRTPYGGQRYSPLTQVNVVNSQLKEAMR--YQTGVKLPDVGERTTYQVTPKIGNTL 225
OY 92 YLANPGDVIQDAIDAKTGDLIWEHRRLPNI 121
DB 226 YICTPHMAIADAATGKEKKYD---PNV 252

RESULT 10
A13064
glucose dehydrogenase gcd [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: A13064
R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutayavln, T.; Levy, R.; Li, M.; McClell Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193
A:Accession: A13064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-778 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL44935.1; PID:g17742589; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: gcd
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match
Best Local Similarity 22.0%; Score 148; DB 2; Length 778;
Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

OY 3 PTSLMA-SAGALALAAPAF-----QVTPVDLLANPPA-----GEMI 42
DB 108 PTGLSYSPNAMPPLVLSVASFVAVGYSMAODPHDQAGSLPQELISAPVYGEVPGDDMH 167
OY 43 SYGNGENYRHSPLTQITTEENVGLOLVWARGMPGKY-----QVTPLIHDGVY 91
DB 168 QYGRTPYGGQRYSPLTQVNVVNSQLKEAMR--YQTGVKLPDVGERTTYQVTPKIGNTL 225
OY 92 YLANPGDVIQDAIDAKTGDLIWEHRRLPNI 121
DB 226 YICTPHMAIADAATGKEKKYD---PNV 252

RESULT 11
F83360
glucose dehydrogenase PA2290 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83360
R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; PMID:120437337
A:Accession: F83360
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <STO>
A:Cross-references: GB:AE004654; GB:AE004091; NID:g9948311; PIDN:AG05678.1; GSPDB:GN00187
A:Experimental source: strain PA01
C:Genetics:
A:Gene: gcd; PA2290
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match
Best Local Similarity 20.4%; Score 137.5; DB 2; Length 803;
Matches 44; Conservative 18; Mismatches 57; Indels 47; Gaps 5;

OY 6 LLM-----ASAGALALAAPAFACVPTVDLLANP----- 36
DB 103 LLMLEPFRRLADGAPRGLTAAGVAVVLAAGAAVGSQ-FTNPQIYGRIDRSGMTST 161
OY 37 -----PAGEMISYQNGENYRHSPLTQITTEENVGLOLVWARGMPGKY-----Q 81
DB 162 APAMPQGDQAVGRTEFGDRYSPLQIIPANVGLEAMR--IRGDDPPTADDDPLELTNE 219
OY 82 VTPLIHDGVYKLANPGDVIQDAIDAKTGDLIWEHRRLPNIATLNSF 127
DB 220 NPLKAVNGMLACTASHKVLADPDTGAEIWFDPQIOSPVGKGF 265

RESULT 12
B83399
quinoprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain PA01)

A:Residues: 1-801 <CLE>
 A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712
 A:Experimental source: strain LMD 79.41
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane F
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:39-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91_93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 18.1%; Score 121.5; DB 1; Length 801;
 Best Local Similarity 30.3%; Pred.No. 0.00043;
 Matches 36; Conservative 15; Mismatches 45; Indels 23; Gaps 4;
 QY 20 PAPAQVPTVDLLANPPAGEMISYGQNGENTRHSPLOTITTENVGQLQVWARGMOPGK 79
 Db 152 PETAAVPGVAE-----SDMPAYGRTOAGVRYSPKQINDQNVKDLKVAMT--LRTGD 202
 QY 80 V-----QVTPLIHGVMYLANPGDVIQAIDAKTGDLWEHRRQLPNIATLNSF 127
 Db 203 LKTNDSGFTTNOVPIKIGNMFFICTAHQOLAIIDPATGKEKW---RFDPKRLKTDKSF 258

Search completed: May 24, 2002, 10:06:04
 Job time: 288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: May 24, 2002, 10:15:48 ; Search time 27.86 Seconds
(without alignments)
177.893 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128

Perfect score: 673
Sequence: 1 MKPTSLIMASAGALALAAP.....DLIWEHRQLPNIATLNSRG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	206	30.6	757	1	DHET_GLUOX
2	183	27.2	738	1	DHET_ACEPO
3	182	27.0	739	1	DHET_ACEU
4	178	26.4	742	1	DHET_ACEAC
5	155	23.0	796	1	DHG_ECOLI
6	137	20.4	623	1	EXAA_PSEAE
7	134	19.9	808	1	DHG_GLUOX
8	121.5	18.1	801	1	DHGA_AICCA
9	112.5	16.7	631	1	DHMI_PARDE
10	111.5	16.6	626	1	DHMI_METOR
11	101.5	15.1	626	1	DHMI_METEX
12	94.5	14.0	809	1	QUIA_AICCA
13	93	13.8	600	1	XOXF_PARDE
14	90	13.4	790	1	QUIA_XANCI
15	78.5	11.7	148	1	CUG8_LOCOMI
16	77	11.4	733	1	PSAB_OOSI
17	76	11.3	233	1	TNEA_MARMO
18	73	10.8	571	1	DHMI_METME
19	73	10.8	776	1	PSAB_HEETR
20	70.5	10.5	540	1	PPBI_RAT
21	70	10.4	441	1	PUN2_THIEFU
22	70	10.4	753	1	YBHJ_ECOLI
23	69.5	10.3	400	1	DACC_ECOLI
24	69	10.2	353	1	INVI_CAPAN
25	68.5	10.2	417	1	OM51_HABIN
26	68.5	10.2	524	1	HEM2_SELMA
27	68.5	10.2	417	1	P60_LISIV
28	68	10.1	159	1	MP63_MCTCU
29	68	10.1	347	1	YAB9_MYCPN
30	68	10.1	661	1	PSAB_PROHO
31	68	10.1	1515	1	GUTB_AZOBRA
32	67.5	10.0	146	1	AZUP_ALCOFA
33	67.5	10.0	261	1	TPIS_MYCLE

34	67.5	10.0	734	1	PSAB_CYACA
35	67.5	10.0	974	1	PHS2_SOLTU
36	67	10.0	331	1	DIV_ECOLI
37	67	10.0	485	1	BIND_STREN
38	67	10.0	500	1	YFJ1_YEAST
39	67	10.0	864	1	STR2_PIG
40	66.5	9.9	1550	1	GLTB_STN3
41	66	9.8	346	1	GPDA_XYLFA
42	66	9.8	369	1	YDHR_ECOLI
43	66	9.8	527	1	YF00_MYCPN
44	66	9.8	755	1	TRZ1_AGRVI
45	65.5	9.7	207	1	DSBA_YERPE

ALIGNMENTS

RESULT ID	1	DHET_GLUOX	STANDARD:	PRT:	757 AA.
AC	005542:				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).				
GN	ADHA.				
OS	Glucobacter oxydans (Glucobacter suboxydans).				
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae; Glucobacter.				
OX	NCBI_TaxID=442;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.				
RC	STRAIN-IFO 12528;				
RX	MEDLINE=97208225; PubMed=9055427;				
RA	Kondo K., Horinouchi S.;				
RT	Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Glucobacter suboxydans and their expression in Acetobacter pasteurianus.;				
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).				
RL	Appl. Environ. Microbiol. 63:1131-1138(1997).				
CC	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.				
CC	-1- COFACTOR: PO4 AND HEME.				
CC	-1- SUBUNIT: HETEROPOLYMER (DEHYDROGENASE AND PROTEIN ADS). THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: D86375; BAA19753.1; -				
DR	HSSP: Q924U7; 1FTG.				
DR	InterPro: IPR001479; Bac_POQ.				
DR	InterPro: IPR002372; Bac_POQ_repeat.				
DR	InterPro: IPR000345; Cyclic_heme_bind.				
DR	Pfam: PF01011; Bacterial_POQ_6.				
DR	PROSITE: PS00363; BACTERIAL_POQ_1; 1.				
DR	PROSITE: PS00364; BACTERIAL_POQ_2; 1.				
DR	PROSITE: PS00190; CYTOCHROME_C_1.				
KW	Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.				
FT	SIGNAL	1	34		
FT	CHAIN	35	757		ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT	MOD RES	35	35		PYRROLIDONE CARBOXYLIC ACID.
FT	ACT_SITE	342	342		BASE (POTENTIAL).

O95166 cyanidium c
P53535 solanum tub
P15286 escherichia
P23118 strongyloce
P43601 saccharomyc
O02799 sus scrofa
P55037 synechocyst
O95677 xylella fas
P75707 escherichia
P75287 mycoplasma
P25017 agrobacteri
O95627 yersinia pe

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 757 AA; 82968 MW; 3989F90E3B947581 CRC64;

Query Match 30.6%; Score 206; DB 1; Length 757;
 Best Local Similarity 43.1%; Pred. No. 1.3e-12;
 Matches 53; Conservative 19; Mismatches 43; Indels 8; Gaps 5;

QY 7 LMSAGALLIAA--PAFQV---TPVDELLANPPAGWISYQNGQENYRHSPLQITTT 61
 Db 16 LLSGAAALFAFAVPAVPAEDPTGTAITSSDNGGHP-GDWLSYGRSYSEORSPDQINT 74
 QY 62 ENVGOLQVWARGMOPGKQV-TPLIHGVMTLANPQVIAQDAKGLIMEHRQLP- 119
 Db 75 ENVGKLKLAHMYDLDTNRGEGTPLIVNVMATITNMSMKRALDAATGKLMSYDPKVP 134
 QY 120 NIA 122
 Db 135 NIA 137

RESULT 2

DHET_ACEPO STANDARD; PRT; 738 AA.
 AC P28036; MEDLINE-91159482; Pubmed-2001402;

DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Acetobacter.
 OX NCBI_Taxid=439;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC STRAIN=NH1028;
 RX MEDLINE-91159482; Pubmed-2001402;
 RA Tanaka T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RT Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes";
 RT Biochim. Biophys. Acta 1088:292-300(1991).
 RL

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: D00635; BAA00528.1; -
 CC PIR: S14270; S14270.
 CC HSSP: Q9Z4J7; 1FT6.

DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT BINDING 650 650 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 27.2%; Score 183; DB 1; Length 738;
 Best Local Similarity 33.3%; Pred. No. 2.3e-10;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALLIAAFAVQVTPVDE-----LLANPPAGWISYQNGQENYRHSPLQITTT 61
 Db 17 TACTICALLISGYATMTASADDGCGATGEATIHADHDHGNNMTGGRYSQDRSPDQINR 76
 QY 62 ENVGOLQVWARGMOPGKQV-TPLIHGVMTLANPQVIAQDAKGLIMEHRQLP- 119
 Db 77 SNGNKLKLAHMYDLDTNRGEGTPLIVNVMATITNMSMKRALDAATGKLMSYDPKVP 136
 QY 120 NIA 122
 Db 137 NIA 139

RESULT 3

DHET_ACEU STANDARD; PRT; 739 AA.
 AC Q44002; 007952;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconacetobacter.
 OX NCBI_Taxid=33995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RA Thurner C.A.K.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: X82894; CAA58066.1; -
 CC DR EMBL: Y09480; CAA70688.1; -
 CC HSSP: Q9Z4J7; 1FT6.

DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RP TOPOLOGY.
 RA MEDLINE-93286127; PubMed-8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RX Cozler G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia coli* modelled on that of methanol dehydrogenase from *Methylobacterium extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC -----
 DR EMBL: X51323; CAA35706.1; -;
 DR EMBL: D12651; BAA02174.1; -;
 DR EMBL: D26562; CAB20298.1; -;
 DR EMBL: AE000122; AAC7335.1; -;
 DR PIR: JVO107; JVO107.
 DR HSSP: P38539; 4AAH.
 DR EcoGene: EG10369; gcd.
 DR InterPro: IPR002372; Bac_POO.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic; Complete proteome.
 FT DOMAIN 1 10
 FT TRANSMEM 11 37
 FT DOMAIN 38 40
 FT TRANSMEM 41 58
 FT DOMAIN 59 62
 FT TRANSMEM 63 81
 FT DOMAIN 82 95
 FT TRANSMEM 96 110
 FT DOMAIN 111 118
 FT TRANSMEM 119 141
 FT DOMAIN 142 796
 FT ACT_SITE 466 466
 FT CONFLICT 59 59
 FT CONFLICT 149 156
 FT CONFLICT 193 193
 FT CONFLICT 666 666
 SQ SEQUENCE 796 AA; 86747 MW; D9HDC705A12894E9 CRC64;

Query Match 23.0%; Score 155; DB 1; Length 796;
 Best Local Similarity 31.9%; Fred. No. 1.4e-07;

Matches 44; Conservative 20; Mismatches 46; Indels 28; Gaps 6;
 QY 6 LIMS-----AGALALAPAPAPVPEVDLLANPAGEWISYCONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTSLADTAPAA-ISPVDQ-----DWAYRNDGQRFSLKQ 165
 QY 59 ITTENYGLQLVMA-----KGMQPGKV--QVPLIHGVMYLANPGVIAQIDAKTGD 109
 DB 186 INADVHLKAWVFRGDKVQPNPDEITNEVTPIKVGDPLYLCTAHQRLFLDAASGK 245
 QY 110 LIWEHRRQLPNATLNSF 127
 DB 246 EKMYHD---PELKTNESEF 260
 RESULT 6
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC 0924J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC *Pseudomonas*.
 OX NCBI_TaxID=287;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a homodimer: sequence of the gene and deduced structural properties of the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardis K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RL [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schobert M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of the genes encoding cytochrome c550 and an adjacent acetaldehyde dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RL [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-20202376; PubMed-10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W., Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POO AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: A1009858; CAA08896.1; -;
 CC EMBL: AE004624; AAG03570.1; -;
 CC EMBL: AF068264; AAC79657.1; -;
 CC PDB: 1FL6; 30-AUG-00.
 CC InterPro: IPR001479; Bac_PQO.
 CC InterPro: IPR002372; Bac_PQO_repeat.
 CC Pfam: PF01011; Bacterial_PQO_7.
 CC PROSITE: PS00363; BACTERIAL_PQO_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_PQO_2; FALSE_NEG.
 CC Oxidoreductase; PQO; Periplasmic; Signal; Calcium; 3D-structure;
 CC Complete proteome.
 CC FT SIGNAL 1 34
 CC FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 CC FT DISULFID 139 140
 CC FT SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;
 SQ
 Query Match 20.4%; Score 137; DB 1; Length 623;
 Best Local Similarity 31.5%; Pred. No. 6.1e-06;
 Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;
 QY 3 PTLNLSAGALAL--LAAPAFQVTPVDELLAN--PPAGEMISYGNQENYRHSPLT 57
 Db 9 PAGLRRLSHCLAFVAVALGSAALAKDVWEDIANDKDTGTVLQGMGTHTHQMSPLK 68
 QY 58 QITENNGQLQVWARM---QPGKVQVPLIHGVMYLANGVYQALDATGDLIME 113
 Db 69 QVNADNVEFKLTPMWSYFSGDEKORQ-ESQAIYSDGVYVYASYRFLADARTGRKLT 127
 QY 114 HRRLPLN 120
 Db 128 YNRRLPD 134
 RESULT 7
 DHG_GLUOX STANDARD; PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconobacter.
 CC NCBL_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.",
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO.

CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 CC PIR: S17716; QPKEX.
 CC InterPro: IPR001479; Bac_PQO.
 CC InterPro: IPR002372; Bac_PQO_repeat.
 CC Pfam: PF01011; Bacterial_PQO_7.
 CC PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 CC Oxidoreductase; PQO; Transmembrane; Periplasmic; Signal.
 CC FT SIGNAL 1 33
 CC FT CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 CC FT TRANSMEM 35 54 POTENTIAL.
 CC FT TRANSMEM 59 76 POTENTIAL.
 CC FT TRANSMEM 94 108 POTENTIAL.
 CC FT TRANSMEM 123 138 POTENTIAL.
 CC FT ACT_SITE 470 470 BASE (POTENTIAL).
 CC FT VARIANT 788 788 H -> N (IN P2 FORM).
 CC FT SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;
 SQ
 Query Match 19.9%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 1.6e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;
 QY 13 ALALLAAPAFQVTPVDELLA-----NPPAGEMISYGNQENYRHSPLTQIT 61
 Db 128 AVIALFASLFTDPHDISGLPTQIANASPADPDNPVASEMHAHYGRTQAGDRSPINQINA 187
 QY 62 ENVGQLQVW-----ARGNPGKV--QVTPPLIHGVMYLANPQDVYQALDAKGDILW 112
 Db 188 TTVSNLKYAMHITKDMNNSNDPGQTEATPIEFNNMLYMSLHQKLFADVAGARGNVK 247
 QY 113 EHRRL 118
 Db 248 VYDPKL 253
 RESULT 8
 DHGA_ACTICA STANDARD; PRT; 801 AA.
 AC P03463;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDHA.
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 CC NCBL_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMD 79.41;
 RX MEDLINE=88289368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 RT dehydrogenase from Acinetobacter calcoaceticus.",
 RL Nucleic Acids Res. 16:6228-6228(1988).

```

CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucose-1,5-lactone +
CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
CC DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X07235; CAA30222.1; -.
CC PIR: S00943; S00943.
CC InterPro: IPR001479; Bac_POQ.
CC Pfam: PF01011; Bacterial_POQ_1; 1.
CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
CC SIGNAL
CC CHAIN 1 33
CC FT TRANSMEM 39 55
CC FT TRANSMEM 59 79
CC FT TRANSMEM 94 108
CC FT TRANSMEM 119 138
CC FT ACT_SITE 471 471
CC SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;

Query Match 18.1%; Score 121.5; DB 1; Length 801;
Best Local Similarity 30.3%; Pred. No. 0.00027;
Matches 36; Conservative 15; Mismatches 45; Indels 23; Gaps 4;

QY 20 PAFAYTVPTDELANPAGEMISYGCNENYRHSPLQTITENVGLOLYMARGMPGK 79
DB 152 PETAQAVPGVAE-----SWPAGRTQAGVRSPLKQINDVKNKDKLKAWT--LRGD 202
QY 80 V-----QVPLIHDSGVLYANPGDVIQAIDAKTGDLIMEHRRLPNIATINSF 127
DB 203 LKTDNDSETTNGVTPITIGNNMFTCAHQDLIAIDPATGKEK---RFDPKLTKDSF 258

RESULT 9
DHM1_PARDE STANDARD; PRT; 631 AA.
ID DHM1_PARDE
AC P12293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RX MEDLINE=87307969; PubMed=3114231;
RA Hams N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
RT structural gene from Paracoccus denitrificans.";
RT J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

```

```

CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL. (IN P-DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17339; AAA8366.1; -.
CC HSP: P38539; 4AAH.
CC InterPro: IPR001479; Bac_POQ.
CC Pfam: PF01011; Bacterial_POQ_1; 1.
CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC SIGNAL
CC CHAIN 1 32
CC FT DISULFID 135 136
CC FT DISULFID 418 447
CC FT ACT_SITE 335 335
CC SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;

Query Match 16.7%; Score 112.5; DB 1; Length 631;
Best Local Similarity 32.6%; Pred. No. 0.0016;
Matches 42; Conservative 25; Mismatches 47; Indels 15; Gaps 9;

QY 4 TSLMASAGALLL-AAFAQVTPYDEL--LANPAGEMISYGCNENYRHSPLQTT 60
DB 12 SSLAMVAMGLAVLTAPATA----NDQLVELAKDPA-NWWTGHDYNAONYSEDTDN 65
QY 61 TENVGLOLYW--ARGMPGKVPTPLINDGVYLANP-GDVIQAIDA-KTGDLLMEHR- 115
DB 66 KENVKQLRPMKSTGVGLHCH-EGTPLYVGDRFHTPTPTTFALDLNPKKILMNKP 124
QY 116 RQLPNIAATL 124
DB 125 KQNPARTAV 133

RESULT 10
DHM1_METOR STANDARD; PRT; 626 AA.
ID DHM1_METOR
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RX STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
RX MEDLINE=89008094; PubMed=2459109;
RA Machin S.M., Hanson R.S.;
RT "Nucleotide sequence and transcriptional start site of the
RT Methylobacterium organophilum XX methanol dehydrogenase structural
RT gene.";
RT J. Bacteriol. 170:4739-4747(1988).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

```

```

CC      reduced acceptor.
CC      -1- COFACTOR: POQ.
CC      -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC      OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC      -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC      MEMBRANE.
CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M22629; AAAH.
CC      HSSP; P38539; AAAH.
CC      DR      InterPro: IPR001479; Bac_POQ.
CC      DR      InterPro: IPR002372; Bac_POQ_repeat.
CC      DR      Pfam; PF01011; Bacterial_POQ_7.
CC      DR      PROSITE; PS00363; BACTERIAL_POQ_1; 1.
CC      DR      PROSITE; PS00364; BACTERIAL_POQ_2; 1.
CC      KW      Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC      FT      SIGNAL      1      28
CC      FT      CHAIN      29      626      METHANOL DEHYDROGENASE SUBUNIT 1.
CC      FT      DISULFID      130      131      BY SIMILARITY.
CC      FT      DISULFID      413      442      BY SIMILARITY.
CC      FT      ACT_SITE      330      330      BASE (POTENTIAL).
CC      SQ      SEQUENCE      626 AA; 68677 MW; 8768F6B8371E5DFE CRC64;

Query Match      16.6%; Score 111.5; DB 1; Length 626;
Best Local Similarity 23.2%; Pred. No. 0.002; Mismatches 38; Indels 65; Gaps 7;
Matches 39; Conservative 26;

QY      10      SAGALALLA-AP-AFAQYPTVDELANPAGE-WISYGOENRHSPLTQITTEYNGQ 66
DB      7      SYSALAMALAPALSSAVANDKVELSKSDNMVWGKNDNNYSSELKQVNSVKQ 66
QY      67      LQLVW-ARGMQPGKVQVTPPLIHGQVWY-----LANGDV----- 99
DB      67      LRPAMTFSTGLNGH-EGAPLVVDGKMYHTSPNNTFALDDDPGHLIMODKRPONPA 125
QY      100      -----IQALDAKTGDLIWE 113
DB      126      RAVACCDLVNRLGATWPGDGKTPALILKTQLDNRVVALNAETGETVWK 173

RESULT 11
DHML METEX STANDARD; PRT; 626 AA.
AC      P16027;
DC      01-APR-1990 (rel. 14, Created)
DT      01-APR-1990 (rel. 14, Last sequence update)
DT      15-JUL-1999 (rel. 38, Last annotation update)
DE      Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE      alpha subunit) (MEDH).
GN      MOXF.
OS      Methylobacterium extorquens.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Methylobacterium group; Methylobacterium.
OX      NCBI_TaxID=408;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AM1 / NCIB 9133;
RC      MEDLINE=90337342; PubMed=2116368;
RA      Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
RT      "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
RT      metx genes involved in methanol oxidation.";
RL      Gene 90:173-176(1990).
RN      [2]
RP      SEQUENCE OF 28-53.

```

```

RC      STRAIN=AM1 / NCIB 9133;
RX      MEDLINE=89350892; PubMed=2504152;
RA      Nunn D.N., Day D., Anthony C.;
RT      "The second subunit of methanol dehydrogenase of Methylobacterium
RT      extorquens AM1.";
RL      Biochem. J. 260:857-862(1989).
RN      [3]
RP      DISULFIDE BONDS.
RX      MEDLINE=95384759; PubMed=7656012;
RA      Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
RT      "The active site of methanol dehydrogenase contains a disulphide
RT      bridge between adjacent cysteine residues.";
RL      Nat. Struct. Biol. 1:102-105(1994).
RN      [4]
RP      X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RX      MEDLINE=95253818; PubMed=7735834;
RA      Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
RT      "The refined structure of the quinoprotein methanol dehydrogenase
RT      from Methylobacterium extorquens at 1.94 A.";
RL      Structure 3:177-187(1995).
CC      -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC      reduced acceptor.
CC      -1- COFACTOR: POQ.
CC      -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC      OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC      -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC      MEMBRANE.
CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M31108; AAA25380.1; -.
CC      DR      PIR; S07908; S07908.
CC      DR      PIR; J00706; J00706.
CC      DR      HSSP; P38539; AAAH.
CC      DR      InterPro: IPR001479; Bac_POQ.
CC      DR      InterPro: IPR002372; Bac_POQ_repeat.
CC      DR      Pfam; PF01011; Bacterial_POQ_7.
CC      DR      PROSITE; PS00363; BACTERIAL_POQ_1; 1.
CC      DR      PROSITE; PS00364; BACTERIAL_POQ_2; 1.
CC      KW      Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC      FT      SIGNAL      1      27
CC      FT      CHAIN      28      626      METHANOL DEHYDROGENASE SUBUNIT 1.
CC      FT      DISULFID      130      131
CC      FT      DISULFID      413      442      BASE (POTENTIAL).
CC      FT      ACT_SITE      330      330
CC      SQ      SEQUENCE      626 AA; 68434 MW; 6498BD0AFD2AD34C CRC64;

Query Match      15.1%; Score 101.5; DB 1; Length 626;
Best Local Similarity 23.2%; Pred. No. 0.019; Mismatches 40; Indels 65; Gaps 7;
Matches 39; Conservative 24;

QY      10      SAGALALLA-AP-AFAQYPTVDELANPAGE-WISYGOENRHSPLTQITTEYNGQ 66
DB      7      SYSALAMALAPALSSAVANDKVELSKSDNMVWGKNDNNYSSELKQVNSVKQ 66
QY      67      LQLVW-ARGMQPGKVQVTPPLIHGQVWY-----LANGDV----- 99
DB      67      LRPAMTFSTGLNGH-EGAPLVVDGKMYHTSPNNTFALDDDPGHLIMODKRPONPA 125
QY      100      -----IQALDAKTGDLIWE 113
DB      126      RAVACCDLVNRLGATWPGDGKTPALILKTQLDNRVVALNAETGETVWK 173

RESULT 12

```

Q10A.ACICA STANDARD; PRT: 809 AA.

AC 059086; Q10A.ACICA

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).

OS Q10A.

GN Acinetobacter calcoaceticus.

OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;

OX NCBI_TaxID=471;

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC STRAIN-BD413 / ADP1;

RX MEDLINE=95095936; PubMed=8002591;

RA Elsmore D.A., Ornstion L.N.;

RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus contains quia, the structural gene for quinate-shikimate dehydrogenase.";

RL J. Bacteriol. 176:7659-7666(1994).

RN [2]

RP SEQUENCE OF 1-20 FROM N.A.

RC STRAIN-BD413 / ADP1;

RX MEDLINE=96011389; PubMed=7592351;

RA Elsmore D.A., Ornstion L.N.;

RT "Unusual ancestry of dehydratases associated with quinate catabolism in Acinetobacter calcoaceticus.";

RL J. Bacteriol. 177:5971-5978(1995).

CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.

CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone -> 5-dehydroquinate + reduced pyrroloquinoline-quinone.

CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.

CC -1- COFACTOR: POO.

CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY, FIRST STEP. THIS PATHWAY ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO PYRROLOQUINOLINE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE PATHWAY.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL, L05770; AAC37161.1; -

DR InterPro: IPR001479; Bac_POO.

DR Pfam: PF01011; Bacterial_POO_7.

DR PROSITE: PS00363; BACTERIAL_POO_1; FALSE_NEG.

DR PROSITE: PS00364; BACTERIAL_POO_2; FALSE_NEG.

KW Oxidoreductase; POO; Quinate metabolism; Transmembrane.

FT TRANSMEM 14 34 POTENTIAL.

FT TRANSMEM 41 61 POTENTIAL.

FT TRANSMEM 68 88 POTENTIAL.

FT TRANSMEM 90 110 POTENTIAL.

FT TRANSMEM 127 147 POTENTIAL.

SO SEQUENCE 809 AA; 88196 MW; 71f67CBEA62BFCB CRC64;

Query Match 14.0%; Score 94.5; DB 1; Length 809;

Best Local Similarity 29.2%; Pred. No. 0.12; Indels 23; Gaps 4;

Matches 35; Conservative 13; Mismatches 49; Indels 23; Gaps 4;

QY 3 PISLWASAGALLAA-PAFAQVTPVTDLLANPAGMISYGOENYRHSPLTQIT 61

DB 148 PHEVYKAGEELPLVPDPARKV-----NMDHGNDAAGSRFVALQINR 193

QY 62 ENVGQQLVW-----ARGMQPG-KVQVPELHDGVMYLANPDVIOAIDATGDLWE 113

DB 194 NNVSXKEMRFRRTGDFTTGTGNGCAEDOMPLOVGNVFLCTPTNNNFALDADSGKQLWK 253

RESULT 13

XOXF-PARDE STANDARD; PRT: 600 AA.

ID XOXF-PARDE

AC P29968;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).

GN XOXF.

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OX Paracoccus.

NCBI_TaxID=266;

RN [1]

RP SEQUENCE FROM N.A.

RA Harms N.;

RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE OF 380-600 FROM N.A.

RC STRAIN-PD 1235;

RX MEDLINE=92041583; PubMed=1657873;

RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F., Stouthamer A.H.;

RT "Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c551 of Paracoccus denitrificans and characterization of the mutant strain.";

RL J. Bacteriol. 173:6971-6979(1991).

CC -1- CORFACTOR: POO (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL, U34346; AAC44555.1; -

DR EMBL, M75583; AAA25574.1; -

DR PIR, A41378; A41378.

DR HSP, P38359; AAAH.

DR InterPro: IPR002372; Bac_POO_repeat.

DR Pfam: PF01011; Bacterial_POO_7.

KW Oxidoreductase; POO; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.

FT ACT_SITE 318 318 BASE (POTENTIAL).

SO SEQUENCE 600 AA; 65159 MW; DCA996F1BC5A3CE CRC64;

Query Match 13.8%; Score 93; DB 1; Length 600;

Best Local Similarity 27.9%; Pred. No. 0.12; Indels 22; Gaps 7;

Matches 36; Conservative 23; Mismatches 48; Indels 22; Gaps 7;

QY 10 SAGALALAAFAQVTPVTDLLANP-----PAGEWISYGOENYRHSPLTQITTE 62

DB 6 NCACIALMSGIAA-----LANEQRAGRDKAOPMAIOMGDYANTRTSTLDQINKD 56

QY 63 NVGQQLVW--ARGMQPGKVQVTPVTLHDGVMYLANP-GDVIOAIDAK-TGDLWEHR-RQ 117

DB 57 NVKDLRVAMFTSGVLRGH-EGSPLVIGDVYVHTPPNRFVAFALDNDCKITLRYEPQ 115

QY 118 LPNIALNS 126

DB 116 DPNVIAVMS 124

Fri May 24 11:27:17 2002

us-08-934-506a-5_copy_1_128.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:30 ; Search time 93.93 Seconds
(without alignments)
235.743 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128
Sequence: 1 MKPTSLMAGALALAAV.....DLIWEHRQLPNIATLNSFG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organellar:*
9: sp-phage:*
10: sp-plant:*
11: sp-protent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	38.8	608	2	093RE9
2	190	28.2	742	2	053362
3	178.5	26.5	691	2	09AR95
4	174	25.9	623	2	09AGW3
5	168.5	25.0	698	2	09KH03
6	167	24.8	786	2	P95466
7	165.5	24.6	708	2	046444
8	156	23.2	695	2	09P9U2
9	155	23.0	785	16	09ARKF6
10	146.5	21.8	777	16	09ZRB3
11	137.5	20.4	803	16	091115
12	134.5	20.0	601	2	09EYW8
13	134.5	20.0	601	16	09ZMY9
14	121	18.0	790	2	09XZS5
15	118.5	17.6	695	2	0934G0
16	112	16.6	601	2	P71509

17	110.5	16.4	629	2	09AQ48	09AQ48 methylolact
18	103	15.3	633	2	024759	024759 hypomicrob
19	93.5	13.9	639	2	P77931	P77931 pseudomonas
20	93.5	13.9	644	2	052551	052551 pseudomonas
21	89	13.2	499	2	051842	051842 thiolactillu
22	83.5	12.4	2832	2	093N05	093N05 agrobacteri
23	83	12.3	359	2	09L935	09L935 methylovoru
24	82.5	12.3	485	17	096XP7	096XP7 sulfolobus
25	82	12.2	668	17	09RRH3	09RRH3 deinococcus
26	81	12.0	613	17	028873	028873 archaeoglob
27	80.5	12.0	647	16	09HUB1	09HUB1 pseudomonas
28	78.5	11.7	509	16	P6368	P6368 mycobacteri
29	77	11.4	846	2	086092	086092 methylovoru
30	76	11.3	306	16	09RT72	09RT72 deinococcus
31	76	11.3	386	16	09RTW8	09RTW8 vibrio chol
32	76	11.3	573	2	059540	059540 methylophil
33	75.5	11.2	407	16	09PG38	09PG38 xylella fas
34	75.5	11.2	684	2	09S282	09S282 streptomyce
35	75.5	11.2	3972	16	P73139	P73139 synecocyst
36	75	11.1	680	5	09N9G7	09N9G7 phallusia m
37	74	11.0	380	16	09HXJ7	09HXJ7 pseudomonas
38	74	11.0	525	16	0931E9	0931E9 staphylococ
39	74	11.0	755	2	09MMA1	09MMA1 agrobacteri
40	74	11.0	969	16	09RZ15	09RZ15 deinococcus
41	73.5	10.9	183	16	09CKL1	09CKL1 pasteurella
42	73.5	10.9	1229	5	09VM05	09VM05 drosophila
43	73	10.8	355	16	0926F4	0926F4 rhizobium m
44	73	10.8	407	16	09AE62	09AE62 caulobacter
45	73	10.8	439	10	004099	004099 brassica ca

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT: 608 AA.
AC 093RE9
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharotogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shibata T., Saito Y.;
RT "Alcohol dehydrogenase.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1;
SQ SEQUENCE 608 AA; 65101 MW; 0ACEC97AE1BA570 CRC64;

Query Match 38.8%; Score 261; DB 2; Length 608;

Best Local Similarity 41.0%; Pred. No. 1.6e-17;
Matches 57; Conservative 20; Mismatches 42; Indels 20; Gaps 2;

QY 10 SAGATALLAAPFAO-----VTPTDELANPAGWISYSONEN 50
DB 16 STALLASLGPFAHDNAAAEPSKAGOSAIENOPTYADDLACKNPANWPLTNGNG 75
QY 51 YRHSPLQTITTEVGOQLVWARGMQPGKVQYTPLIHGVMTLANPGVYIAIDAKTGL 110
DB 76 WGSPLDQINKDNVGLQLVMSRTMEPGSNEGAALAYNVIFLGNTWITQALIDKTS 135
QY 111 IWEHRQLPNIAT-LSFG 128
DB 136 IWEYRRKLPASAKFINSIG 154

RESULT 2
 ID 053362 PRELIMINARY; PRT; 742 AA.
 AC 053362; 044159;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
 OS Acetobacter pasteurianus (Acetobacter turbidans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 AC Acetobacter.
 OX NCBI_TaxID=438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NC11380.
 RX MEDLINE-94042848; PubMed-8226628;
 RA Takemura H., Kondo K., Hironouchi S., Beppu T.;
 RT Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
 RT pasteurianus.";
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL; D13893; BAA40252.1; -.
 DR HSSP; 0924J7; 1FLG.
 DR InterPro; IPR001479; Bac_PQO.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 6.
 DR PROSITE; PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
 SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 28.2%; Score 190; DB 2; Length 742;
 Best Local Similarity 38.8%; Pred. No. 2.2e-10;
 Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGAALLAAPAAQAVPTVDL--ANPPAGEMISYQNGQENYRHSPLQTITTEWGLQ 68
 DB 23 AALPAAPAAQAVPTVDL--ANPPAGEMISYQNGQENYRHSPLQTITTEWGLQ 82
 QY 69 LWARGMOPGKVO-VTPLIHDGVMTLANPGVIAQIDAKTGDLWEHRRQLP-NIA 122
 DB 83 LAMVYTLDTNRQGEATPLVVDGIMVATNWSKMLADATGKLQMDPKVPGNIA 138

RESULT 3
 ID 09AF95 PRELIMINARY; PRT; 691 AA.
 AC 09AF95;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF355798; AKK27220.2; -.
 SQ SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 26.5%; Score 178.5; DB 2; Length 691;
 Best Local Similarity 37.8%; Pred. No. 2.8e-09;
 Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

QY 39 GEMISYQNGQENYRHSPLQTITTEWGLQDLMWARGMOPGK-VQVTPLIHDGVMTLANPG 97
 DB 41 GEMVTHGVDAAGTRYSPLAQITTPNNAKELGLWMSYDLESNGVATITVVDGVMYVIAFW 100

QY 98 DVIAQIDAKTGDLWEHRRQLP 119
 DB 101 SVVHALDVRSKRLMTYDPEVP 122

RESULT 4
 ID 09AGW3 PRELIMINARY; PRT; 623 AA.
 AC 09AGW3;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=66174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
 RT "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF326086; AAK15506.1; -.
 DR HSSP; 0924J7; 1FLG.
 DR InterPro; IPR002372; Bac_PQO_repeat.
 DR Pfam; PF01011; Bacterial_PQO; 4.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 25.9%; Score 174; DB 2; Length 623;
 Best Local Similarity 38.0%; Pred. No. 6.8e-09;
 Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLWASGALALAAPAAQAVPTVDL--ANPPAGEMISYQNGQENYRHSPLQTIT 59
 DB 7 KPTSLWASGALALAAPAAQAVPTVDL--ANPPAGEMISYQNGQENYRHSPLQTIT 65
 QY 60 TTEWVGLQDLMWARGMOPGKVO-VTPLIHDGVMTLANPGVIAQIDAKTGDL 111
 DB 66 NTIDVNAVLPAMSEFSGEKQRGQ--AGV--LVHDGVYIYNAAS.SRIFAIDARSGRL 120
 QY 112 WEHRRQLP 120
 DB 121 WEYNAARLPD 129

RESULT 5
 ID 09KH03 PRELIMINARY; PRT; 698 AA.
 AC 09KH03;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE TERAHYDROFURFYL ALCOHOL DEHYDROGENASE.
 OS Alcaligenes eutrophus (Ralstonia eutropha).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21122557; PubMed-11222593;
 RA Zarrin G., Schrader T., Andreesen J.R.;
 RT "Catalytic and Molecular Properties of the Quinolomoprotein
 RT tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha
 RT Strain Bo."
 RL J. Bacteriol. 183:1954-1960(2001).
 DR EMBL; AF277373; AAF86335.1; -.
 DR HSSP; 0924J7; 1FLG.

CC -|- FUNCTION: CATALYZES THE DIRECT OXIDATION OF ALIPHATIC ALCOHOLS
CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC
CC -|- COFACTOR: PQQ, HEME, AND CALCIUM.

CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C; 1.
 DR Signal: PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
 FT TYPE 1.
 FT BINDING 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match 24.6%; Score 165.5; DB 2; Length 708;
 Best Local Similarity 31.8%; Pred. No. 5.6e-08;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PSLMASGALALLAFAFAQVTPVT-----DELLANPP-AGEWISYQONEN 50
 DB 12 PGRWVLAACG--SAAAFQGTGRNAAAVRVDDGFITANAARPPDMFTIGVDIAE 69
 QY 51 YHSPITQTTENVGOLQVMARQPGK-VQVPLIHGVYLANPGDVIQAIIDAKTGD 109
 DB 70 TYSRSDQINANVNDLGLAMSYNLESTRGVATPVVDGIMVYASMSVVAIDITRGN 129
 QY 110 LIMEHROP 118
 DB 130 RIMTYDPQ 138

RESULT 8
 Q9F902 PRELIMINARY; PRT; 695 AA.
 AC Q9F902:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2.";
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAC09249.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 23.2%; Score 156; DB 2; Length 695;
 Best Local Similarity 33.6%; Pred. No. 4.8e-07;

Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;
 QY 15 ALLAPAFQVTPVTDELL--ANPPAGEWISYQONENRHSPLTQTTENVGOLQVMA 72
 DB 16 ALLVA-AGAAKAYDEAIRASEODGSEWLSHGRTVAEORRSPKQIDAGNVGRLGLAWY 74
 QY 73 RGMOPGK-VQVPLIHGVYLANPGDVIQAIIDAKTGDLLWEHROL 118
 DB 75 LDLENKRGLEATPLVSDGLYASLSMSRWMAVDLRSGKRLMQEDPOV 121

RESULT 9
 Q98KF6 PRELIMINARY; PRT; 785 AA.
 AC Q98KF6:
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN ML1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MAF303099.
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002997; BAB48858.1; .
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 KW Complete proteome.
 SQ SEQUENCE 785 AA; 84083 MW; 1DC224B7A1C25470 CRC64;

Query Match 23.0%; Score 155; DB 16; Length 785;
 Best Local Similarity 31.7%; Pred. No. 7e-07;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;

QY 8 WASAG-----ALLAPAFQVTPVTD-----ELLANPAGE 40
 DB 111 WARRGLAGDGRAPLILAVLASIYAGYSMTADPKGIDALDPTKIVPRKANGNDVPAGE 170
 QY 41 WISYQONENRHSPLTQTTENVGOLQVMAQMGPKV-----QVTPLIHNG 89
 DB 171 WHYGRQFGQSRSPLDQITPDVAVMLQPAWT--YFTGVKGPDDIGETTYVTPPLKIGD 228
 QY 90 VMTLANPGDVIQAIIDAKTGDLLWEH 114
 DB 229 TLVTCRPHNPAIYVDAATCKEKRWY 253

RESULT 10
 Q92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYRROLOQUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21368234; PubMed=11474104;
 RA Gailbert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Boutry M., Bowers L., Buhrmester J., Cadieu E., Capela D., Chain P.,
 RA Gloux S., Godrie T., Goffeau A., Goding B., Gouzy J., Gurjal M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kallman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masny D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 RA Ranspiger U., Surzycki R., Thebaud P., Vandenbol M.,
 RA Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont *Sinorhizobium meliloti*."
 RL Science 293:668-672(2001).
 DR EMBL: AL591785; CAC45553.1;
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 777 AA; 84174 MW; D2F478A38B10430 CRC64;

Query Match 21.8%; Score 146.5; DB 16; Length 777;
 Best Local Similarity 38.5%; Pred. No. 4.8e-06;
 Matches 35; Conservative 9; Mismatches 34; Indels 13; Gaps 2;

QY 35 NPPAGEMISYGNQENYRHSPLTQITTEVNGOLOLVNARGMOPGV-----QVT 83
 DB 159 NPPAGEMHOGYRTPRGQRTSPLODRPENVMALQVAMQ--YTGDKVLPEDVSETTYQVT 216
 QY 84 PLIHGCVMTLANPGDVIAIDAKTGDLWEH 114
 DB 217 PLKVKNTLVYCTPHNMALDAATGQEKWKY 247

RESULT 11
 Q91115 PRELIMINARY; PRT; 803 AA.
 AC Q91115;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-OCT-2001 (TREMREL. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GCD OR PA2290.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Binkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Hickey M.J., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larij K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004654; AAG05678.1;
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 KW Complete proteome.
 SQ SEQUENCE 803 AA; 86217 MW; 7DDF46B0FE5ED249 CRC64;

Query Match 20.4%; Score 137.5; DB 16; Length 803;
 Best Local Similarity 26.5%; Pred. No. 3.9e-05;
 Matches 44; Conservative 18; Mismatches 57; Indels 47; Gaps 5;
 QY 6 LLM-----ASGALALALAPAFAPQVTPVDELILANP----- 36

DB 103 LLMPEFRREPLADGAPPLCTAALGAVVLAGAANGSQ-FTRPGQIVGRIDRDSGMTST 161
 QY 37 -----PAGEWISYGNQENYRHSPLTQITTEVNGOLOLVNARGMOPGV-----Q 81
 DB 162 APAMPDGDWQAYGRTERGDSYSPKQITPANVQGLEARN--IRTDGLPLADPLELTNE 219
 QY 82 VTPLIHGCVMTLANPGDVIAIDAKTGDLWEHRRQLPINATINSF 127
 DB 220 NTPKVGNGMYLACTANSHKVALPDVGAIEIRDPDQIOSPVGKGF 265

RESULT 12
 Q92WY8 PRELIMINARY; PRT; 601 AA.
 AC Q92WY8;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
 GN MXAF.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RM1021;
 RA Fennel B.J., Tiwari R.P., Dilworth M.J.;
 RT "Regulation of C1 assimilation in *Sinorhizobium meliloti*."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF309488; AAG31643.1;
 DR HSSP; P38539; 4AAH.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 SQ SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 20.8%; Score 134.5; DB 2; Length 601;
 Best Local Similarity 33.0%; Pred. No. 5.4e-05;
 Matches 38; Conservative 27; Mismatches 41; Indels 9; Gaps 7;

QY 14 LALLAAPAFAPQVTPVDEL--LANPAGEWISYGNQENYRHSPLTQITTEVNGOLOLVN 71
 DB 8 LAIISGGAGVAPRANDELQKIIDP--NQMALQYGDVIANLKYSLDQJNKNVQKLGAVAW 66
 QY 72 --ARCMOPGVYVPLIHGCVMTLANP-GDVIQALD-AKTGDLWEHR-RQLPIN 121
 DB 67 TESTGVLRGH-EGSPLVIGDLMYVTPPTVYALDLSKDOIWKYKPPKODPNV 120

RESULT 13
 Q92WY9 PRELIMINARY; PRT; 601 AA.
 AC Q92WY9;
 DT 01-DEC-2001 (TREMREL. 19, Created)
 DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
 DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
 DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC
 1.1.99.8).
 GN SMB20173.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid psymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Goding B., Puehler A.;
 RT "The complete sequence of the 1,683-kb psymb megaplasmid from the N2-

RT fixing endosymbiont *Sinorhizobium meliloti*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603642; CAC48573.1; -
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424EFED15ADBE CRC64;

Query Match 20.0%; Score 134.5; DB 16; Length 601;
 Best Local Similarity 33.0%; Pred. No. 5.4e-05;
 Matches 38; Conservative 27; Mismatches 41; Indels 9; Gaps 7;

QY 14 LALLAPAFQVTPVDEL--LANPPAGEISTGQNGENYRHSPLQTITTEVGOLOLW 71
 DB 8 LAIWSGGGAQVAFANDELQKLDLP-NOWAIQTGDYANLRYSKLDQINKDNVGLQVAM 66
 QY 72 --ARGNQPGKVQVTPLIHGVWYLANP-GDVIAID-AKTGLIMEHR-ROLPI 121
 DB 67 TFEVGLRKH-EESPLVIGDLKAVHFPFNTYALDLSKQGVWKEPKDPPV 120

RESULT 14

O9X2S5 PRELIMINARY; PRT; 790 AA.

AC O9X2S5
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GDHB.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.
 OX NCBI_TaxID=53336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1056R;
 RX MEDLINE=99303331; PubMed=10376838;
 RA Pujol C.J., Kado C.I.;
 RT "gdhB, a gene encoding a second quinoprotein glucose dehydrogenase in
 RT Pantoea citrea, is required for pink disease of pineapple."
 RL Microbiology 145:1217-1226(1999).
 DR EMBL; AF050503; AAD3735.1; -
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
 DR SEQUENCE 790 AA; 85681 MW; 88BB7625E0A5F089 CRC64;

Query Match 18.0%; Score 121; DB 2; Length 790;

Best Local Similarity 32.6%; Pred. No. 0.0017;
 Matches 31; Conservative 13; Mismatches 39; Indels 12; Gaps 3;

QY 27 PVTDELANPPAGEISTGQNGENYRHSPLQTITTEVGOLOLW-----RGMQPK 79
 DB 152 PATQPVSAKG---DWTAVGCTQAGDRYSLSQINESNVKNLQVAMTYQSGDVKRPPDPGE 208

QY 80 V--QVTPLIHGVWYLANPGDVIAIDAKTGDLIW 112
 DB 209 ITNESTPLKGNLITCTAHQILVAIDAATGCKEK 243

RESULT 15

O934G0 PRELIMINARY; PRT; 695 AA.

AC O934G0
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE LUPANINE HYDROXYLASE PRECURSOR.
 GN LQH.
 OS Pseudomonas sp. DH2001.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.
 OX NCBI_TaxID=163360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
 RA Rogozinski J.;
 RT "Cloning, sequencing and analysis of the gene for lupanine
 RT hydroxylase, a quinoxycytocrome c from a Pseudomonas sp."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ318095; CAC67410.1; -
 KW Signal.
 FT SIGNAL
 FT CHAIN
 SQ SEQUENCE 695 AA; 74956 MW; 9096C6387E457FE0 CRC64;

QY 38 AGEWISTGQNGENYRHSPLQTITTEVGOLOLVARGMQPGKVQV-TPLIHGVWYLANP 96
 DB 36 SGNWSLLGGNEQHYFSLADVKNKSNVKNLGSFTDMEAGDGLVGNPLVADGVITYQGP 95
 QY 97 GDVIAIDAKTGDLIMEH 114
 DB 96 PGKIYANDLKTGKNLMTY 113

Query Match 17.6%; Score 118.5; DB 2; Length 695;
 Best Local Similarity 35.9%; Pred. No. 0.0025;
 Matches 28; Conservative 8; Mismatches 41; Indels 1; Gaps 1;

Search completed: May 24, 2002, 10:17:31
 Job time: 744 sec

Fri May 24 11:27:18 2002

us-08-934-506a-5_copy_1_128.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:01 ; Search time 117.28 Seconds
(without alignments)
127.856 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135
Perfect score: 710
Sequence: 1 MKPTSLIMASAGALLLAAP.....ROLPNITATLNSEFGEPTRGMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	710	100.0	579	19 AAM37873	Alcohol and/or ald
2	640	90.1	579	19 AAM37874	Alcohol and/or ald
3	587.5	82.7	578	19 AAM37875	Alcohol and/or ald
4	567	79.9	579	19 AAM37876	Alcohol and/or ald
5	193.5	27.3	754	21 AAB35987	Sorbitol dehydrog
6	187	26.3	738	12 AAR13993	A.altoecetigenes me
7	187	26.3	738	13 AAR20192	ADH complex protei
8	178	25.1	742	11 AAR05235	Amino acid sequenc
9	158.5	22.3	443	22 AAG24430	Novel human diagen
10	145.5	20.5	740	20 AAM95019	Sorbitol dehydrog
11	129	18.2	948	22 AAG25909	Novel human diagen

12	129	18.2	1510	22 AAG21573	Novel human diagen
13	77.5	10.9	353	22 AAB47447	MOAP F5. Haemophi
14	76	10.7	824	18 AAM09614	Pathogenic Staphy1
15	75.5	10.6	1419	22 AAB59208	Drosophila melanog
16	74	10.4	282	22 AAB81927	Acromonium cellulo
17	74	10.4	294	21 AAB35811	Chitinase amino ac
18	74	10.4	224	21 AAB13001	Chitinase protein
19	73.5	10.4	1229	22 AAB71736	Drosophila melanog
20	72.5	10.2	343	19 AAM055001	Bacillus thermoloe
21	72.5	10.2	462	17 AAM05395	Human SH3P17 prote
22	71	10.0	526	22 AAB99895	Xanthomonas sp. ch
23	70.5	9.9	171	22 AAB32330	Novel human secret
24	70.5	9.9	475	22 AAG92298	C glutamicum prote
25	70.5	9.9	843	20 AAB34485	Porphoromonas ging
26	70.5	9.9	849	20 AAB34360	Porphoromonas ging
27	70.5	9.9	5435	22 AAE10145	Streptomyces nous
28	70	9.9	1016	22 AAB34349	Staphylococcus aur
29	69.5	9.8	275	21 AAG42781	Arabidopsis thalia
30	69.5	9.8	1029	22 AAG26194	Novel human diagen
31	69.5	9.8	1029	22 AAG28658	Novel human diagen
32	69	9.7	711	22 AAB69743	Drosophila melanog
33	69	9.7	957	22 AAB62401	Drosophila melanog
34	68.5	9.6	384	22 AAB85255	Plant porpobillio
35	68.5	9.6	551	22 AAB85251	Plant thioedoxin-
36	68.5	9.6	551	22 AAB85252	Myobacterium tube
37	68	9.6	159	19 AAM63034	Myobacterium tube
38	68	9.6	159	21 AAB14325	Myobacterium tube
39	68	9.6	424	22 AAB59716	Drosophila melanog
40	68	9.6	718	22 AAG91907	C glutamicum prote
41	67.5	9.5	275	21 AAG16656	Arabidopsis thalia
42	67.5	9.5	735	22 AAG24896	Novel human diagen
43	67.5	9.5	851	14 AAB41333	113 kd ISF-3alpha
44	67.5	9.5	851	16 AAR2077	Recognition factor
45	67.5	9.5	851	17 AAM03166	Human STAT2. Homo

ALIGNMENTS

RESULT 1	
AAW37873	AAW37873 standard; Protein: 579 AA.
ID	AAW37873
XX	
AC	AAW37873:
XX	
DT	10-AUG-1998 (first entry)
XX	
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.
XX	
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KW	2-keto-L-gulonidic acid; L-ascorbic; inhibition.
XX	
OS	Glucanobacter oxydans.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	/note="signal peptide"
FT	24..579
FT	/note="mature protein"
XX	
PN	EP832974-A2.
XX	
PD	01-APR-1998.
XX	
PF	11-SEP-1997; 97EP-0115801.
XX	
PR	19-SEP-1996; 96EP-0115001.
XX	
PA	(HOEF) HOFFMANN LA ROCHE & CO AG F.
XX	
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX	

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1: Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA:

Query Match 100.0%; Score 710; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 1,1e-69;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPTSLMASAGALIALAPAFAYPTVDELLANPAGENISYSGONENYRHSPLTOIT 60
 |||
 Db 1 mptslmasagalaialapafayptvdelianpagsewisysqnenyrrhsplqit 60
 QY 61 TENVGQLQVWARGMQPKQVTPPLIHGVMYLANPGDVIAQIDAKTGDLWEHRRQLPN 120
 |||
 Db 61 tenvgqlqvwarqmkgkvtpplihdgvmylanpgdviaqidaiktgdliwehrrqlpn 120
 QY 121 IATLNSFGPEPTRGMA 135
 |||
 Db 121 iatlntfgeptrgma 135

RESULT 2

AAW37874
 ID AAW37874 standard; Protein: 579 AA.

XX AAW37874;

DT 10-AUG-1998 (first entry)

XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KM 2-keto-L-gulonic acid; L-ascorbic; inhibition.

XX Gluconobacter oxydans.

XX Key Location/Qualifiers

FT Peptide 1..23 /note="signal peptide"

FT Protein 24..579 /note="mature protein"

XX EP832974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1: Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA:

Query Match 90.1%; Score 640; DB 19; Length 579;
 Best Local Similarity 91.9%; Pred. No. 6,1e-62;
 Matches 124; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKPTSLMASAGALIALAPAFAYPTVDELLANPAGENISYSGONENYRHSPLTOIT 60
 |||
 Db 1 mktsslmasagalaialapafayptvdelianpagsewisysqnenyrrhsplqit 60
 QY 61 TENVGQLQVWARGMQPKQVTPPLIHGVMYLANPGDVIAQIDAKTGDLWEHRRQLPN 120
 |||
 Db 61 tenvgqlqvwarqmkgkvtpplihdgvmylanpgdviaqidaiktgdliwehrrqlpn 120
 QY 121 IATLNSFGPEPTRGMA 135
 |||
 Db 121 iatlntfgeptrgma 135

RESULT 3

AAW37875
 ID AAW37875 standard; Protein: 578 AA.

XX AAW37875;

DT 10-AUG-1998 (first entry)

XX Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.

XX Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KM 2-keto-L-gulonic acid; L-ascorbic; inhibition.

XX Gluconobacter oxydans.

XX Key Location/Qualifiers

FT Peptide 1..23 /note="signal peptide"

FT Protein 24..578 /note="mature protein"

XX EP832974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29053.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 578 AA;

Query Match 82.7%; Score 587.5; DB 19; Length 578;
 Best Local Similarity 83.0%; Pred. No. 3.8e-56;
 Matches 112; Conservative 9; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MKPTSLMASAGALALLAPAFQVTVTDDELLANPAGEMISYQNGENYRHSPLTOIT 60
 Db 1 mklttllqssaaallvltipalag-taltdemlanpagewinyngngenyhsplqt 59
 QY 61 TENVGQLQVWARGMQPKVQVTPPLIHGVMYIANGDVIOAIDAKTGLIWEHRRQLPN 120
 Db 60 adnvgqlqlvwarameagkivtpldvnyvnylanpgdvlgaldagqllwehrrqlpn 119
 QY 121 IATLNSFGEPTRGMA 135
 Db 120 laltnstfgeptrgma 134

RESULT 4
 AAW37876
 ID AAW37876 standard; Protein; 579 AA.
 AC AAW37876;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note="signal peptide"
 FT Protein 24..579
 FT /note="mature protein"
 XX
 PN EP832974-A2.
 XX
 PD 01-APR-1998.
 XX
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX
 PT

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29054.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 44-46; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 79.9%; Score 567; DB 19; Length 579;
 Best Local Similarity 78.5%; Pred. No. 7e-54;
 Matches 106; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
 QY 1 MKPTSLMASAGALALLAPAFQVTVTDDELLANPAGEMISYQNGENYRHSPLTOIT 60
 Db 1 mptlllrtsaavllltapaafqvtvptidellanpagewinyngngenyhsplqt 60
 QY 61 TENVGQLQVWARGMQPKVQVTPPLIHGVMYIANGDVIOAIDAKTGLIWEHRRQLPN 120
 Db 61 adnvgqlqlvwarameagavqtpmhdgvmylanpgdvlgaldagqllwehrrqlpa 120
 QY 121 IATLNSFGEPTRGMA 135
 Db 121 valtnagqdkrkyva 135

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 XX
 DT 01-MAR-2001 (first entry)
 XX
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulonic acid.
 XX
 OS Gluconobacter oxydans.
 OS
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-1B00736.
 XX
 PR 23-APR-1999; 99WO-1B00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI: 2000-687351/67.
 DR N-PSDB: AAC83153.
 XX
 PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative

CC See also AAQ20383-84, and -86-88.
 XX
 SQ Sequence 738 AA;

Query Match 26.3%; Score 187; DB 13; Length 738;
 Best Local Similarity 32.4%; Pred. No. 9.1e-12;
 Matches 44; Conservative 23; Mismatches 59; Indels 10; Gaps 3;

QY 10 SAGLALLAPAFQAVPTVDE-----LLANPPAGEMISGONQENYRHSPLTQITT 61
 Db 17 tagtcaallsgyatmasaddggatgcatgealhhadhpnmwlygrtysdqrysplqgnr 76
 QY 62 ENVGQLQVWARGMQPGKVQV-TPLIHGVMTLANPGDVLOAIDAKTGDLWEHRROP- 119
 Db 77 srygnlklaylaldtnrgegtplvldgymyatlmwmmkavdaatgklwsgdprvpg 136
 QY 120 NIATLNSFGEPTRGMA 135
 Db 137 niadkgcctlvnrgaa 152

RESULT 8

AAQ05235 standard; protein; 742 AA.

AAQ05235;

04-AUG-1990 (first entry)

Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases 1-2229 of alcohol dehydrogenase (ADH) gene.

Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;

Acetobacter acetii K1006 (FERM-7528); plasmid pAA721.

Acetobacter acetii strain K1006 (FERM-7528).

JP02000452-A.

05-JAN-1990.

30-OCT-1987; 87JP-0273190.

30-OCT-1987; 87JP-0273190, JP-075069.

(QPPP) QP CORP (QPPF).

WPI; 1990-047990/07.

N-PSDB; AAQ91811.

Cloning DNA, plasmid and microbe contg. it -
 PT contains alcohol dehydrogenase gene derived from Acetobacter

acetii K1006 strain

Disclosure; Fig 4; 8pp; Japanese.

Also new are a recombinant plasmid contg. its encoding DNA, and a

microorganism transformed with the plasmid.

A DNA fragment was obtd. from A. acetii

strain K1006 and was transferred to an ADH-defective strain. The

resulting plasmid, pAA721, contg. its gene was inserted into A.

pasteureus strain IFO 3191 by the triparent method. Its gene is

useful for improving Acetobacter culturing.

Sequence 742 AA;

Query Match 25.1%; Score 178; DB 11; Length 742;

Best Local Similarity 36.2%; Pred. No. 9.1e-11;

Matches 46; Conservative 19; Mismatches 58; Indels 4; Gaps 3;

11 AGALALLAFAFAQVTVTDELL--ANPPAGEMISGONQENYRHSPLTQITTENVGQLQ 68

Db 23 aalpyaavparaagqntgealhhadhpnmwlygrtysdqrysplqgnrsvqdlk 82

QY 69 IWMARGMQPGKVQV-TPLIHGVMTLANPGDVLOAIDAKTGDLWEHRROP-NIATLNS 126

Db 83 llygtldtnrgeatplvldgymyatlmwmmkavdaatgklwsgdprvpgniadkyc 142

QY 127 FGEPTRG 133

Db 143 cdtlvnrg 149

RESULT 9

ABG24430 standard; Protein; 443 AA.

ABG24430;

18-FEB-2002 (first entry)

Novel human diagnostic protein #24421.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HSE-) HXSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS8617.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity

Claim 20; SEQ ID No 54789; 103pp; English.

The invention relates to isolated polynucleotide (I) and

polypeptide (II) sequences. (I) is useful as hybridization probes,

polymerase chain reaction (PCR) primers, oligomers, and for chromosome

and gene mapping, and in recombinant production of (II). The

polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (I) is useful in gene therapy techniques

to restore normal activity of (II) or to treat disease states involving

(II). (II) is useful for generating antibodies against it, detecting or

quantitating a polypeptide in tissue, as molecular weight markers and as

a food supplement. (II) and its binding partners are useful in medical

imaging of sites expressing (II). (I) and (II) are useful for treating

disorders involving aberrant protein expression or biological activity.

The polypeptide and polynucleotide sequences have applications in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity

and to produce other types of data and products dependent on DNA and

amino acid sequences. ABG00010-ABG30377 represent novel human

diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 443 AA;

Query Match

22.3%; Score 158.5; DB 22; Length 443;

Best Local Similarity 32.0%; Pred. No. 6.4e-09;

Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;

OY 6 LMAS-----AGALALLAPAFQVPTVDELLANPPAGMISYCONENYRHSPLTQ 58
 DB 168 ltwagfndpqlngtclsdarphae-ispvqd-----dwpaygrngsgqrftspkq 218
 OY 59 ITTENVGOLQLVMA-----RGMPGRV--QVTPLIHDGVWYLANPGDVYIOAIDAKTGD 109
 DB 219 lhadvnhlkeawvfrtgdkpndgplnevpikvgtllylctahgrlfaaldaasgk 278
 OY 110 LIWEHRQLPNATLNSFGEPT-RGMA 135
 DB 279 ekwhyd---pelktnesifqhvrcyvs 302

RESULT 10

AAM95019
ID AAM95019 standard; Protein: 740 AA.

XX AC AAM95019;

DT 21-MAY-1999 (first entry)

DE Sorbitol dehydrogenase (SLDH) protein sequence.

XX Sorbitol dehydrogenase; SLDH; open reading frame; ORF2; L-sorbose;

KM D-sorbitol; Vitamin C; enzyme.

XX OS Gluconobacter suboxydans.

FH Key Location/Qualifiers

FT Peptide 1..24

FT /note= "signal sequence"

FT Protein 25..740

FT /note= "mature protein"

XX EP897984-A2.

XX PD 24-FEB-1999.

XX PF 13-AUG-1998; 98EP-0115231.

XX PR 21-AUG-1997; 97EP-0114432.

XX PA (HOEF) HOFEMANN LA ROCHE & CO AG F.

XX PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

XX DR WPI: 1999-134646/12.

XX DR N-PSDB; AAX21501.

PT New D-sorbitol dehydrogenase gene and recombinant protein - useful
 for production of L-sorbose, an intermediate in vitamin C production

XX PS Claim 1; Fig 3A-D; 39pp; English.

XX This represents a sorbitol dehydrogenase (SLDH) protein. The DNA
 CC encoding the SLDH enzyme also encodes an open reading frame (ORF2)
 CC product upstream of the SLDH open reading frame, needed for SLDH activity
 CC in vivo. Host cells transformed by a vector comprising the SLDH DNA
 CC sequence are used for the recombinant expression of the sorbitol
 CC dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 CC producing L-sorbose from D-sorbitol. L-sorbose is an important
 CC intermediate in vitamin C production.

XX Sequence 740 AA;

Query Match

20.5%; Score 145.5; DB 20; Length 740;

Best Local Similarity 34.0%; Pred. No. 3.5e-07;

Matches 34; Conservative 25; Mismatches 30; Indels 11; Gaps 5;

OY 23 AQVTPVDELLANPPAGMISYCONENYRHSPLTQITTEENVGOLQ--VMARGM--QPG 78
 DB 86 sqvpmagpqsanpargdwaygrddhqrtyrslsetpnaasklvafyhtgsyprg 145
 OY 79 KV-----QVTPLI-HDGVWYLANPGDVYIOAIDAKTGDLIW 112
 DB 146 qvnkwaettpkvgdgllyrcsamndlk-ldpatgkqlw 184

RESULT 11

ABG25909
ID ABG25909 standard; Protein: 948 AA.

XX AC ABG25909;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25900.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS90096.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX PS Claim 20; SEQ ID No 56268; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 948 AA;

PF treat Haemophilus influenzae infection or associated disease, e.g.,
 PF otitis media and conjunctivitis -
 XX
 PS Disclosure: Fig 1: 29pp: English.
 XX
 CC This sequence represents the major outer membrane protein p5 of
 CC non-typeable H. influenzae. One or more surface exposed loops of this
 CC protein may be replaced with a modified peptide of the invention. Each
 CC of these peptides contain an LRI(f) peptide which is a 19 amino acid
 CC peptide derived from the sequence of MOMP p5 from strain nH1128,
 CC representing amino acids Arg117 to Gly135. This peptide represents the
 CC third exposed loop of p5 and is a potential B cell epitope. The loops
 CC of the invention are modified in terms of being in a non-native
 CC environment in the recombinant outer membrane protein. The modified
 CC MOMP p5 may be used to induce an immune response in a mammal to
 CC prevent or treat Haemophilus influenzae infection or associated
 CC disease, e.g., otitis media, sinusitis, conjunctivitis, or lower
 CC respiratory tract infection.
 CC
 SQ Sequence 353 AA;

Query Match 10.9%; Score 77.5; DB 22; Length 353;
 Best Local Similarity 24.4%; Pred. No. 4.1;
 Matches 39; Conservative 21; Mismatches 71; Indels 29; Gaps 5;

OY 1 MKPTSLWASGALALLAPAFQVTPYTDLELANPAGF-----WISYGO 46
 DB 1 mkktalavag-----laasvagaapgentfyagvkagafidglalareykvgyhr 56
 OY 47 NOENRHSPLQITTE-----VGQLVWARGMPGKQVTPPLHHDGVMTLANPQDV 99
 DB 57 nsftgyvfgyqqlngnmlglavelyddfggrakgrevkvtvthnhtlslksyev 116
 OY 100 IQAIPD-AKTGDLI--WEHRQLEPNIAIATLNSFGEPTGMA 135
 DB 117 legldvygkavalrvsdgklynnensclkklghehtrara 156

RESULT 14

AAW09614
 ID AAW09614 standard; Protein: 824 AA.

AC AAW09614;
 XX
 DF 20-MAY-1997 (first entry)
 XX
 DE Pathogenicity: vaccine; immunoassay; antibody.
 XX
 KW Pathogenicity; vaccine; immunoassay; antibody.
 XX
 OS Staphylococcus aureus strain NCTC 8325-4 (ISP8).
 XX
 PM W09641878-A1.
 XX
 PD 27-DEC-1996.
 XX
 PF 13-JUN-1996; 96WO-AU00353.
 XX
 PR 13-JUN-1995; 95AU-0003507.
 XX
 PA (AUSU) UNITV AUSTRALIAN NAT.
 XX
 PI Adham El-Adhami W, Matthei KI, Stewart PR;
 XX
 DR WPI: 1997-065465/06.
 DR N-PSDB; AAT47757.
 XX
 PT Nucleic acids specific to pathogenic Staphylococcus aureus strains -
 XX
 PS Claim 30; Page 62-65; 88pp: English.
 XX

CC An 824-amino acid protein (AAW09614) and 192-amino acid protein
 CC (AAW09615) were identified as products of 2 open reading frames in
 CC a nucleic acid (AAT47757) isolated from pathogenic Staphylococcus
 CC aureus strain ISP8. The 2 gene products are highly hydrophilic
 CC and have high surface probability and antigenic index measurements,
 CC suggesting that they are surface expressed. The polypeptides,
 CC their fragments, and antibodies raised against them can be used in
 CC vaccines and as immunoassay reagents for detecting pathogenic S.
 CC aureus strains.
 CC
 SQ Sequence 824 AA;

Query Match 10.7%; Score 76; DB 18; Length 824;
 Best Local Similarity 24.8%; Pred. No. 19;
 Matches 33; Conservative 18; Mismatches 34; Indels 48; Gaps 7;

OY 19 APAPFAQVT-----PYTDLELANPAGFEMISYGOENRHSPLQITTEWGOQLIV 70
 DB 683 apgtckvregqgkxklttptlknpltgeliskgskettkdpineit----- 732
 OY 71 WARGMPGKQVTPPLHD-----GVMTLANP--GDVIO-AIDAKT----- 107
 DB 733 -----eygpetltpnrhdefdpxlpcgkeevykpjknpcgvdvrvpvdsvtkygpv 787
 OY 108 -GDVMEHRROLP 119
 DB 788 kgdsive-keelp 799

RESULT 15

ABB59208
 ID ABB59208 standard; Protein: 1419 AA.

AC ABB59208;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 4416.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL03311.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 4416; 21pp + Sequence Listing: English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

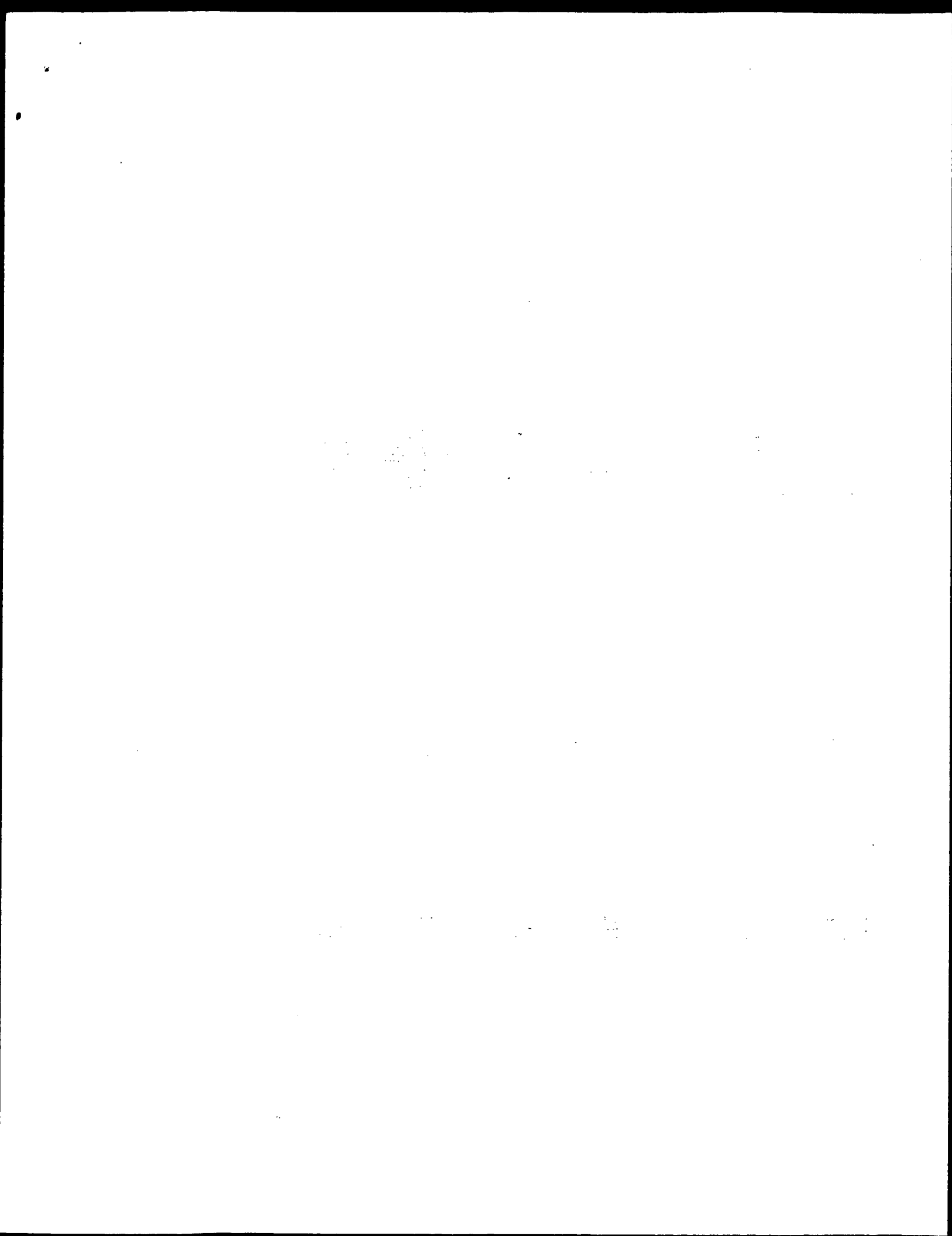
CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB7072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 1419 AA;

Query Match 10.6%; Score 75.5; DB 22; Length 1419;
 Best Local Similarity 22.6%; Pred. No. 46;
 Matches 30; Conservative 15; Mismatches 65; Indels 23; Gaps 3;

QY 9 ASAGALALLAPFAQVTPVTDELIANPAGEWISYGN-----QENYRHSPLTQITTE 62
 Db 716 asrtatlaaaaaaagltptatvaaaatppaasakafnlndvssfsstnkrhgatssnmr 775
 QY 63 NVGQLQVWARGMOPKRVQVTPLIHDSVWYLANPGDVIQAIDAKTGDLIWEHRRQLPNTA 122
 Db 776 ssnqaaaaaaatkaaatatvp-----aapngllscrksns-----qnsplya 818
 QY 123 TLNSGEPTRGMA 135
 Db 819 tpxltltpttrfia 831

Search completed: May 24, 2002, 10:05:02
 Job time: 241 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:02:57 ; Search time 44.35 Seconds

(without alignments)
74.351 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 710
Sequence: 1 MKPISLMAAGALALAAP.....ROLNPIATLNSFGEPTKMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193.5	27.3	754	4	US-09-296-284-4
2	187	26.3	738	1	US-07-985-458-3
3	184.5	26.0	720	4	US-09-296-284-25
4	145.5	20.5	739	3	US-09-136-251-2
5	72.5	10.2	462	4	US-08-630-915A-38
6	70.5	9.9	540	1	US-08-368-071-11
7	70.5	9.9	540	1	US-08-458-181-11
8	70.5	9.9	540	5	PCT-US93-02172-11
9	68	9.6	159	3	US-08-796-792-2
10	67.5	9.5	851	1	US-08-369-796-2
11	67.5	9.5	851	2	US-08-852-091-2
12	67.5	9.5	851	2	US-08-820-754-2
13	67.5	9.5	851	3	US-08-956-652-2
14	67.5	9.5	851	3	US-08-956-869-2
15	67.5	9.5	851	3	US-09-012-710-2
16	67.5	9.5	851	3	US-08-948-547-2
17	67.5	9.5	851	4	US-09-087-465-4
18	67.5	9.5	851	4	US-09-364-970-2
19	67.5	9.5	851	4	US-09-556-273-2
20	67.5	9.5	851	4	US-08-956-653A-2
21	67.5	9.5	851	5	PCT-US95-17025-2
22	67.5	9.5	852	1	US-08-276-099A-13
23	67.5	9.5	852	1	US-08-781-890-13
24	67.5	9.5	974	2	US-08-868-786-6
25	66.5	9.4	210	4	US-09-247-155-117
26	65.5	9.2	876	2	US-08-510-215A-2
27	65.5	9.2	1229	3	US-09-310-293-2

28	65.5	9.2	1229	4	US-09-579-376-2	Sequence 2, Appl
29	64.5	9.1	286	4	US-08-964-127-4	Sequence 4, Appl
30	64.5	9.1	286	4	US-09-496-692-4	Sequence 2, Appl
31	64.5	9.1	520	4	US-08-964-127-2	Sequence 2, Appl
32	64.5	9.1	520	4	US-09-496-692-2	Sequence 2, Appl
33	64.5	9.1	523	2	US-08-923-772-2	Sequence 2, Appl
34	64.5	9.1	523	2	US-09-385-287-2	Sequence 2, Appl
35	64	9.0	966	2	US-08-868-786-2	Sequence 2, Appl
36	63.5	8.9	777	2	US-08-874-678-3	Sequence 3, Appl
37	63.5	8.9	777	3	US-08-643-839-3	Sequence 3, Appl
38	63.5	8.9	820	4	US-09-313-677-21	Sequence 2, Appl
39	63.5	8.9	926	4	US-09-313-677-19	Sequence 2, Appl
40	63.5	8.9	933	4	US-09-313-677-17	Sequence 19, Appl
41	63.5	8.9	967	4	US-09-313-677-17	Sequence 17, Appl
42	63.5	8.9	1298	1	US-08-222-616-33	Sequence 33, Appl
43	63.5	8.9	1298	1	US-08-340-011-2	Sequence 2, Appl
44	63.5	8.9	1298	3	US-08-901-710-2	Sequence 2, Appl
45	63.5	8.9	1298	5	PCT-US95-04228-33	Sequence 33, Appl

ALIGNMENTS

```
RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucanobacter Suboxydans Sorbitol Dehydrogenase. Genes
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Glucanobacter suboxydans
US-09-296-284-4

Query Match      27.3%; Score 193.5; DB 4; Length 754;
Best Local Similarity 37.4%; Pred. No. 6.4e-14;
Matches 49; Conservative 17; Mismatches 56; Indels 9; Gaps 3;

QY 14 LALLAAPAAQVTPV-----TDELLANPPAGEWISTGQNOENTRHSPLDITTEWNGQ 66
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 17 LGCNAALACATSPVALADTGTATINADQHPGDMWSYGRYSSEORYSLDITRDNMSN 76

QY 67 LGLVNRAGQPGQVQV--TPLIHDGVWYLYANPGDVIQADIDAKGLIWEHRDLP-NIATL 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 77 LKLAHWYDDITDNNGGCGTPLIVDGVWYATTTNMSKMAIDATGKLUMSDPKVPGNIDR 136

QY 125 NSGFEPTRGMA 135
   | | |
DB 137 GCCDYNRGAA 147

RESULT 2
US-07-985-458-3
; Sequence 3, Application US/07985458
; Patent No. 5344777
; GENERAL INFORMATION:
; APPLICANT: Tamaki, Toshimi;
; APPLICANT: Takemura, Hiroshi;
; APPLICANT: Takayama, Kenji;
; APPLICANT: Fukaya, Masahiro;
; APPLICANT: Okumura, Hajime and
; APPLICANT: Kawamura, Yoshiya
```

Query Match	26.3%;	Score 187;	DB 1;	Length 738;
Best Local Similarity	32.4%;	Pred. No. 3.5e-13;		
Matches	44;	Conservative	23;	Mismatches 59;
			Indels	10;
			Gaps	3

Query Match 26.0%; Score 184.5; DB 4; Length 720;
Best Local Similarity 38.4%; Pred. No. 6.6e-13;
Matches 43; Conservative 18; Mismatches 46; Indels 5; Gaps 3

RESULT 4
 US-09-136-251-2
 : Sequence 2, Application US/09136251A
 : Patent No. 6127156
 :
 : GENERAL INFORMATION:
 : APPLICANT: HOSHINO, Tatsuo
 : APPLICANT: MIYAZAKI, Taro
 : APPLICANT: OJIMA, Setsuko
 : APPLICANT: TOMIYAMA, Masako
 : APPLICANT: TOKIYAMA, No. 6127156ibumi
 : TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
 : FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
 : CURRENT APPLICATION NUMBER: US/09/136,251A
 : CURRENT FILING DATE: 1998-08-19
 : EARLIER APPLICATION NUMBER: EP 97114432.4
 : EARLIER FILING DATE: 1997-08-21
 : NUMBER OF SEQ ID NOS: 9
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 2

```

; ORGANISM: Gluconobacter suboxydans
; FEATURE:
; NAME/KEY: SIGNAL
;

```


COUNTRY: US
ZIP: 10111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,792
FILING DATE: 06-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/011,364
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hone, William J.
REGISTRATION NUMBER: 26,739
REFERENCE/DOCKET NUMBER: 07763/03301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-765-5070
TELEFAX: 212-258-2291
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

FRAGMENT TYPE: Internal
US-08-796-792-2

Query Match

Best Local Similarity 9.6%; Score 68; DB 3; Length 159;
Best Local Similarity 21.7%; Pred. No. 2.3;
Matches 36; Conservative 19; Mismatches 49; Indels 62; Gaps 7;

QY 1 MKPTSLMASAGALALLAPAFQAVPTVDELLANPAGEMWISYGCNENRHSPLTOIT 60

DB 6 MTTAAVAVVMAAIAITFAFVLAALAPITGKL-----GSELTMA-- 43

QY 61 TENYGOLOLWV-----ARGMQPG-----KVQVTPLIH-----DG 89

DB 44 TDIVGQVGLKMKVSLKSTPAVIGPVAGQVWEATATVNAIGSVTPAVSQFNARTADG 103

QY 90 VMT-----LANP-----GDVIGADATGDLIWEHRQLPNIAFLNS 126

DB 104 INRVIMQAAGPPTISGATIPGEGOSTGKITYDVTGPSPTIVAMNN 149

RESULT 10

US-08-369-796-2
Sequence 2, Application US/08369796
Patent No. 5716622

GENERAL INFORMATION:

APPLICANT: James E. Darnell, Jr.

APPLICANT: Zilong Wen

APPLICANT: Curt M. Horvath

APPLICANT: Zhong Zhong

TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/369,796

FILING DATE: 06-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-116

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 851 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-369-796-2

Query Match

Best Local Similarity 9.5%; Score 67.5; DB 1; Length 851;

Best Local Similarity 22.2%; Pred. No. 28;

Matches 37; Conservative 20; Mismatches 59; Indels 51; Gaps 8;

QY 1 MKPTSLMASAGALALLAPAFQAVPTVDELLANPAGEMWISYGCNENRHSPLTOIT 60

DB 459 MNQSLMASAVLWMLSPNLQNO-----QFNSNPKAWSLGLPALSWQSSSYVGRGL 512

QY 47 NOENYRHSPLTOITENYGO-----LOLVMA-----RGMQPGKV-----QVTPLIHDG 89

DB 513 NSDQ-----LSMLRKLKLGQNCRTEDPLLSWADFTKRSPPGKILPEWTWLKILLEVHDH 567

QY 90 VMTLANPQDVIOAIDAKTGDLIWEHRQLPNIAFLNSGCEPTRG 133

DB 568 IKDLWMDGRIMGFVRS-----QERRLLKRTMGTGTFLLRSESESEG 608

RESULT 11

US-08-852-091-2

Sequence 2, Application US/08852091

Patent No. 5883228

GENERAL INFORMATION:

APPLICANT: James E. Darnell, Jr.

APPLICANT: Zilong Wen

APPLICANT: Curt M. Horvath

APPLICANT: Zhong Zhong

TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/852,091

FILING DATE: 06-MAY-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/369,796

FILING DATE: 06-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-116

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 851 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-852-091-2

Query Match

Best Local Similarity 9.5%; Score 67.5; DB 2; Length 851;

Best Local Similarity 22.2%; Pred. No. 28;

Matches 37; Conservative 20; Mismatches 59; Indels 51; Gaps 8;

QY 1 MKPTSLMASAGALALLAPAFQAVPTVDELLANPAGEMWISYGCNENRHSPLTOIT 60

DB 459 MNQSLMASAVLWMLSPNLQNO-----QFNSNPKAWSLGLPALSWQSSSYVGRGL 512

QY 47 NOENYRHSPLTOITENYGO-----LOLVMA-----RGMQPGKV-----QVTPLIHDG 89

DB 513 NSDQ-----LSMLRKLKLGQNCRTEDPLLSWADFTKRSPPGKILPEWTWLKILLEVHDH 567

QY 90 VMTLANPQDVIOAIDAKTGDLIWEHRQLPNIAFLNSGCEPTRG 133

DB 568 IKDLWMDGRIMGFVRS-----QERRLLKRTMGTGTFLLRSESESEG 608

RESULT 12

US-08-820-754-2

Sequence 2, Application US/08820754

Patent No. 5976835

GENERAL INFORMATION:

APPLICANT: Darnell Jr., James E.

APPLICANT: Schindler, Christian W.

APPLICANT: Fu, Xian-Yuan

APPLICANT: Wen, Zilong

TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN

TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESS: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/820,754

FILING DATE: 19-MAR-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,185

FILING DATE: 11-MAR-1994

APPLICATION NUMBER: US 07/980,498

FILING DATE: 23-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/854,296

FILING DATE: 19-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO US93/02569

FILING DATE: 19-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/126,588

FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-073 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEFAX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 851 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-820-754-2

Query Match 9.5%; Score 67.5; DB 2; Length 851;

Best Local Similarity 22.2%; Pred. No. 28;

Matches 37; Conservative 20; Mismatches 59; Indels 51; Gaps 8;

QY 90 VMYLANPDVYIAIDAKTGDLIWEHRROLPNIAI---LNSFGPTRG 133

Db 568 LKDLMDGRIMGFVSRS-----QERRLKKTKMSGFLLRFSSESG 608

RESULT 13

US-08-956-652-2

Sequence 2, Application US/08956652

Patent No. 6013475

GENERAL INFORMATION:

APPLICANT: Darnell Jr., James E.

APPLICANT: Schindler, Christian W.

APPLICANT: Fu, Xian-Yuan

APPLICANT: Wen, Zilong

TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN

TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESS: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,652

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,185

FILING DATE: 11-MAR-1994

APPLICATION NUMBER: US 07/980,498

FILING DATE: 23-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/854,296

FILING DATE: 19-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/126,588

FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-073 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEFAX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 851 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-956-652-2

Query Match 9.5%; Score 67.5; DB 3; Length 851;

Best Local Similarity 22.2%; Pred. No. 28;

Matches 37; Conservative 20; Mismatches 59; Indels 51; Gaps 8;

QY 1 MKPTSLIMASAGALALAAPFAQVTPYTDLLANPPAGW-----ISYQ 46

Db 459 MNOLSTAMASVLMFNLSPLNLO-----QFSPNPKAPWSLGPALSMQFSSYVGRGL 512

QY 1 MKPMSLIMASAGALALLAAPAFQVTPVTDELLANPAGEW-----ISYQ 46

QY 90 VMYL : 1

QY 90 VMYL : 1

Fri May 24 11:27:18 2002

us-08-934-506a-5_copy_1_135.rai

Page 8

Db 568 LKDLMDGRIMGVSR-----QERRLLKKTMSGTFLLRFSSESSEG 608

Search completed: May 24, 2002, 10:02:58
Job time: 232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:04 ; Search time 56.04 Seconds
(without alignments)
231.479 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135
Perfect score: 710
Sequence: 1 MKPTSLMASAGALALAAP.....ROLPNIALINSFGEPTGMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	26.9	742	2 A49340	alcohol dehydrogen
2	187	26.3	738	2 S14270	alcohol dehydrogen
3	179	25.2	742	2 JS0326	alcohol dehydrogen
4	165.5	22.3	708	2 SS2317	quinohemoprotein e
5	158.5	22.3	796	1 JV0107	glucose dehydrogen
6	158.5	22.3	796	2 H90644	glucose dehydrogen
7	158.5	22.3	796	2 H85495	glucose dehydrogen
8	157.5	22.2	796	2 AG0523	glucose dehydrogen
9	148	20.8	778	2 G98221	glucose dehydrogen
10	148	20.8	778	2 A13064	glucose dehydrogen
11	143	20.1	803	2 F83360	glucose dehydrogen
12	140.5	19.8	601	2 E95863	alcohol dehydrogen
13	137	19.3	623	2 B83399	quinoprotein alcoh
14	134	18.9	808	1 OPRKX	glucose dehydrogen
15	123	17.3	801	1 S00943	glucose dehydrogen
16	107	15.1	809	2 B98314	probable quinone d
17	107	15.1	809	2 A12968	hypothetical prote
18	101.5	14.3	626	2 JU0706	alcohol dehydrogen
19	99	13.9	809	2 A55547	quinone-shikimate
20	93.5	12.1	639	2 JC4881	polyvinyl-alcohol
21	86	11.8	613	2 F69424	conserved hypotet
22	83.5	11.8	2802	2 F97686	cyclic beta-(1-2)
23	83.5	11.8	2831	2 A12911	beta (1-->2) gluc
24	82	11.5	668	2 C75264	probable serine/th
25	81.5	11.5	393	2 AD0350	probable lipoprote
26	80.5	11.3	647	2 C83012	probable chemotaxi
27	80.5	11.3	772	2 H86016	hypothetical prote
28	78.5	11.1	148	1 S05638	cuticle protein 8
29	78.5	11.1	509	2 C70624	probable two compo

30 78 11.0 207 2 AH1548 3-methyladenine DN
31 77.5 10.9 353 2 C64187 outer membrane pro
32 77 10.8 733 2 S78376 photosystem I p700
33 76.5 10.8 224 2 AD2115 phosphoribosylform
34 76.5 10.8 400 2 AD0604 D-alanyl-D-alanine
35 76 10.7 306 2 D75562 N-acetyl-gamma-glu
36 76 10.7 386 2 A82284 conserved hypotet
37 76 10.7 573 2 S68591 methanol dehydroge
38 76 10.7 580 2 A80994 gamma-glutamyltran
39 75.5 10.6 407 2 C82804 conserved hypotet
40 75.5 10.6 684 2 T36771 probable integral
41 75.5 10.6 3972 2 S75251 hypothetical prote
42 74.5 10.5 969 2 A75634 MCRB-related prote
43 74 10.4 380 2 C83171 conserved hypotet
44 73.5 10.4 347 2 A83617 probable oxidoredu
45 73.5 10.4 584 2 AH2506 ser/thr protein Ki

ALIGNMENTS

RESULT 1
A49340 alcohol dehydrogenase (EC 1.1.1.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence-revision 07-Apr-1994 #text-change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:DJ1893; NID:9517067; PIDN:BAA40252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 26.9%; Score 191; DB 2; Length 742;
Best local similarity 37.0%; Pred. No. 5.8e-11;
Matches 47; Conservative 18; Mismatches 58; Indels 4; Gaps 3;
QY 11 AGALLAAPAFAPQVPTVDL--ANPPAGEWISGQNGENRHSPLQITTEVGOLO 68
DB 23 AALPFAAVPARADGCGNTGEALIHADDPENWLSYGRYSRQYSPLOIKRNSVNGDLK 82
QY 69 LVNARGMQEGKV-VTPLIHDGVNMLANPGDVLOATDQTLWEHRQLP-NIATLNS 126
DB 83 LAWYTLDTNRGQCATPLVVDGIMVATWMSKMEALDAATGKLWQYDPKVGNIADKGC 142
QY 127 FGEPTRG 133
DB 143 CDTVNRG 149

RESULT 2
S14270 alcohol dehydrogenase, membrane-bound (EC 1.1.1.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
A:Variety: Strain NB11028
C:Date: 21-Nov-1993 #sequence-revision 26-Jul-1996 #text-change 20-Jun-2000
C:Accession: S14270
R:Takemura, H.; Takemura, H.; Takemura, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:9216186
A:Experimental source: strain NB11028


```

A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A>Status: preliminary
A:Molecule type: Protein
A:Residues: 1-20 <YAM>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <FUD>
A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BAA05580.1; PID:g473791
A:Experimental source: strain K-12 substrain W3110
R:Yamada, M.; Asaoaka, S.; Sailer, M.H.; Yamada, Y.
J. Bacteriol. 175, 568-571, 1993
A>Title: Characterization of the gcd gene from Escherichia coli K-12 W3110 and regulatid
A:Reference number: 141228; MUID:93123180
A:Accession: 141228
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148,'H','150','KRCRT','157-192','K','194-665','H','667-796<RES>
A:Cross-references: GB:D12651; NID:g216555; PIDN:BAA02174.1; PID:g216556
C:Genetics:
A:Gene: gcd
A:Map position: 3 min
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase, pyrroloquinoline quinone, respiratory chain; transmembrane F
F:11-37/Domain: transmembrane #status predicted <TM1>
F:41-59/Domain: transmembrane #status predicted <TM2>
F:63-81/Domain: transmembrane #status predicted <TM3>
F:96-110/Domain: transmembrane #status predicted <TM4>
F:120-140/Domain: transmembrane #status predicted <TM5>
F:93_95/Binding site: ubiquinone (Arg, Asp) #status predicted
F:466/Active site: Asp #status predicted

Query Match      22.3%; Score 158.5; DB 1; Length 796;
Best Local Similarity 32.0%; Pred. No. 1.1e-07;
Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;

OY 6 LHMAS-----ACGLALLAPAFQVTPYDELLANPPAGEWISYGNOENYRISPLTQ 58
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 LTMAEFPNDPQETNGTLADATPAEA-ISPVADQ-----DWPAYGRNGEGQRSPSLKQ 185
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 59 ITTENVGOLQIYWVA-----KGMPGKV--GVPELIHDGVYLVLPDVIQAIDAKTG 109
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 186 INADVNVHLKEKWERYRGTDVKPNPDGETNETNPPIIKXGDPTLYLCTANORLFALDAASGK 245
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 110 LIWEHRRLPNTATLNSFGEPY-RGMA 135
    | : | : | | | | | | | | | |
Db 246 EKWHYD---PELKTNESFOHYTCRGVS 269
    | : | : | | | | | | | | | |

RESULT 6
H90644
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 050952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90644
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasanara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:1158796
A:Accession: H90644
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAB33551.1; PID:g133559584; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952

```

C:Genetics: A:Gene: EC60128 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 Query Match 22.3%, Score 158.5; DB 2; Length 796;
 Best Local Similarity 32.0%, Pred. No. 1,1e-07;
 Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;
 QY 6 LTMAS-----AGLALLAPAFQAQVTPVYDELLANPPAGEMISYGQNGENYRHSPLTQ 58
 Db 135 LTMAGFNDPQELNGTILSADATPAEA-ISPVAQD-----DWPYVGRNGQGRSPKQ 185
 59 ITTENVGOLQLVMA-----RGMQPKV--QVTPPLIHGVMYLANPEDVIOAIDAKTGD 109
 QY 186 INADVHNLKEMWVRFGDVKQPNDPGEITNEVPIKVGDTLYLCTAQRFLFALDAASGK 245
 Db 110 LIWEHRQDLPNATLNSFGPEPT-RGMA 135
 QY 246 EKWHYD---PELKTINESFOHYTCRGVS 269
 RESULT 7
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL9333
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perena, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouzis, K.; Apodaca
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1796 <STO>
 A:Cross-references: GB:A8005174; NID:912512839; PIDN:AAG54428.1; GSPTDB:GN00145; UWGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 C:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 Query Match 22.3%, Score 158.5; DB 2; Length 796;
 Best Local Similarity 32.0%, Pred. No. 1,1e-07;
 Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;
 QY 6 LTMAS-----AGLALLAPAFQAQVTPVYDELLANPPAGEMISYGQNGENYRHSPLTQ 58
 Db 135 LTMAGFNDPQELNGTILSADATPAEA-ISPVAQD-----DWPYVGRNGQGRSPKQ 185
 59 ITTENVGOLQLVMA-----RGMQPKV--QVTPPLIHGVMYLANPEDVIOAIDAKTGD 109
 QY 186 INADVHNLKEMWVRFGDVKQPNDPGEITNEVPIKVGDTLYLCTAQRFLFALDAASGK 245
 Db 110 LIWEHRQDLPNATLNSFGPEPT-RGMA 135
 QY 246 EKWHYD---PELKTINESFOHYTCRGVS 269
 RESULT 8
 AG0523
 glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typh
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: This species has also been called Salmonella typh
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0523
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:g16501455; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0191
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 22.2%; Score 157.5; DB 2; Length 796;
 Best Local Similarity 32.7%; Pred. No. 1.4e-07;
 Matches 48; Conservative 21; Mismatches 49; Indels 29; Gaps 7;

QY 6 LIMS-----AGALALAPAFQVTPYTDLLANPAGEMISYQNGENYRHSPLTQ 58
 DB 135 LTMAGFNDPQINSTLRADATPA-ATSSSIAD-----DWPAYGRNCGORYSPLKQ 185
 QY 59 ITTENVGOLQVMA-----RGMOPGV--QVTPLIHDGVMTLANPGDVIOAIDAKTGD 109
 DB 186 ITADNVHOLKAWFERGCDLKQPNDPGEITNEVTPIKVGTLYICTAHQRLFALDAASGK 245
 QY 110 LIWEHRQLPNIATLNSFGEPT-RGMA 135
 DB 246 EKMFD---POLKTDSSFOVHTCRGVS 269

RESULT 9
 glucose dehydrogenase PA2290 [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: G98221
 R:Gooder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wolham, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
 A:Reference number: A97359; PMID:11743194
 A:Accession: G98221
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-778 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK89297.1; PID:g15159131; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L1436
 A:Map position: linear chromosome
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 20.8%; Score 148; DB 2; Length 778;
 Best Local Similarity 30.0%; Pred. No. 1.2e-06;
 Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

QY 3 PTSLMA-SAGALALAPAF-----QVTPYTDLLANP-----GEMT 42
 DB 108 PTGLSTSPNAMPVLVLSFAVAGYSMAQDPHDQAGSLPQELASAPVYGGVDPDGMH 167
 QY 43 SYQNGENYRHSPLTITTEVNGOLQVMAKMGPKV-----QVTPLIHDGV 91
 DB 168 QYGRTPYGGORYSPLTVNVNVSQLEAMR--YQTGDVLPDDVGETTYQVTPPLKIGNTL 225
 QY 92 YLANPGDVIOAIDAKTGDLIWEHRQLPNI 121
 DB 226 YICTPHNMAIADAATGKMKYD---PNV 252

RESULT 10
 AI3064
 glucose dehydrogenase gcd [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AI3064
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moore, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClair, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AI3064
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-778 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AAL44935.1; PID:g17742589; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: gcd
 A:Map position: linear chromosome
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 20.8%; Score 148; DB 2; Length 778;
 Best Local Similarity 30.0%; Pred. No. 1.2e-06;
 Matches 45; Conservative 19; Mismatches 50; Indels 36; Gaps 6;

QY 3 PTSLMA-SAGALALAPAF-----QVTPYTDLLANP-----GEMT 42
 DB 108 PTGLSTSPNAMPVLVLSFAVAGYSMAQDPHDQAGSLPQELASAPVYGGVDPDGMH 167
 QY 43 SYQNGENYRHSPLTITTEVNGOLQVMAKMGPKV-----QVTPLIHDGV 91
 DB 168 QYGRTPYGGORYSPLTVNVNVSQLEAMR--YQTGDVLPDDVGETTYQVTPPLKIGNTL 225
 QY 92 YLANPGDVIOAIDAKTGDLIWEHRQLPNI 121
 DB 226 YICTPHNMAIADAATGKMKYD---PNV 252

RESULT 11
 F83360
 glucose dehydrogenase PA2290 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83360
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: AB2950; MUID:20437337
 A:Accession: F83360
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-803 <STO>
 A:Cross-references: GB:AE004654; GB:AE004091; NID:g9948311; PIDN:AAG05678.1; GSPDB:GN00187
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: gcd; PA2290
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 20.1%; Score 143; DB 2; Length 803;
 Best Local Similarity 26.9%; Pred. No. 4e-06;
 Matches 47; Conservative 20; Mismatches 60; Indels 48; Gaps 6;

QY 6 LLM-----ASGALALAPAFQVTPYTDLLANP----- 36
 DB 103 LLMWFRRLPADGPAIPGLTAALGAAVAGSO-FTNPQOIGRIDRDSGMTST 161
 QY 37 -----PAGEWISYQNGENYRHSPLTITTEVNGOLQVMAKMGPKV-----Q 81
 DB 162 AAPMPDGMQAVGRTGFEGRISPLKQITPANGLLEAMR--IRIGDLPTADPLELTNE 219

RESULT 15

500943

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Acinetobacter calcoaceticus* strain 102
C. Species: *Acinetobacter calcoaceticus*

C;Species: *Acinetobacter calcoaceticus*

```
C:\Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
```

C; Accession: S00943

R; Cleton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
Nucleic Acids Res 15 6338 1988

Nucleic Acids Res. 16, 6228, 1988

A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from *Thiobacillus*

A;Reference number: S00943; MUID:88289368

A;Accession: S00943

A: Molecule type: DNA
A: Accession: 300343

A: molecule type: DNA
A: Residues: 1-801 <CTE>

A;Residues: 1-801 <CLE>
A;Cross-references: EMBL:Y07329F, NTD:Q39711, PDB:Q112033 1, PDB:30713

A; Cross-references: EMBL:X07235; NID:g3

A; Experiment

C;Function:

A;Description: catalyzes the

```

12: description: catalyzes the oxidation of D-glucose to D-glucos
13: Pathway: respiratory chain

```

A, pathway: respiratory chain
C: Superfamily: glucose dehydrogenase (pyrro)quinoline-quinone)

C: Keywords: oxidoreductase; pyrroloquinoline quinone; superfamily: glucose dehydrogenase (pyrroloquinoline

c;keyworas: oxidoreductase; pyrroloquinoline quinone, E:8-35/Domajm; tttttttttttt #attttttttttt smvt,

F;9-35/Domain: transmembrane #status predicted <TM1>

F;39-57/Domain: transmembrane #status predicted <TM2>

F;61-79/Domain: transmembrane #status predicted <TM3>

F;94-108/Domain: transmembrane #status predicted <TM4>

F:118-137/Domain: transmembrane #status predicted <TM5>
1,24 100/Domain: transmembrane #status predicted <TM4>

F:91,93/Binding site: ublquinone (Arg, Asp) #status predicted
F:471/Active site: Asp #status predicted

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

20 PAF4QVTPVTEDELLANP PAGEDWISYQNGQENYRHSPLTQITTEVNGOLVWARGNQPGK 79

№	Вопрос	Правильный ответ	Ваш ответ
153	Решите задачу: в первом классе 15 человек, во втором – 12 человек, в третьем – 10 человек. Сколько всего человек в трех классах?	37	37
154	Решите задачу: в первом классе 15 человек, во втором – 12 человек, в третьем – 10 человек. Сколько всего человек в трех классах?	37	37

QY 80 V-----QVTPLIHDGVMYLANPBGDYIAIDAKTGDLINHRQLPNIAITLNSFG 128

[illegible]

OV 129 EPT-RGM 134

QY 129 EPT-RGM 134

Search completed: May 24, 2002, 10:06:05
Job time: 289 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:49 ; Search time 27.86 Seconds
(without alignments)
187,622 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 710
Sequence: 1 MKPTSLMASAGALALAAP.....ROLPNATLNSFGEPTRGMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	211	29.7	757	1 DHET_GLUOX	005542 gluconobact
2	187	26.3	738	1 DHET_ACERO	P28036 acetobacter
3	186	26.2	739	1 DHET_ACREU	044002 acetobacter
4	179	25.2	742	1 DHET_ACEAC	P18278 acetobacter
5	158.5	22.3	796	1 DHG_ACEOL	P15877 escherichia
6	137	19.3	623	1 EXAA_PSEAE	092477 pseudomonas
7	134	18.9	808	1 DHG_GLUOX	P27175 gluconobact
8	123	17.3	801	1 DHGA_ACECA	P05455 acinetobact
9	119.5	16.8	631	1 DHM1_PARDE	P12283 paracoccus
10	113.5	16.0	626	1 DHM1_METOR	P15279 methyllobact
11	101.5	14.3	626	1 DHM1_METEX	P16027 methyllobact
12	99	13.9	600	1 XOXE_PARDE	P29968 paracoccus
13	99	13.9	809	1 GUN2_ACECA	059086 acetobact
14	91.5	12.9	790	1 QUN1_XANCI	09x78 xanthomonas
15	88.5	11.1	148	1 CU08_LOCM1	P11734 locusta mig
16	78	11.0	207	1 3MGH_LISIN	092089 listeria in
17	77.5	10.9	353	1 OM51_HABIN	P43840 haemophilus
18	77	10.8	733	1 PSAB_ODOSI	P49480 odontella s
19	76	10.7	233	1 TNFA_MARMO	035734 marmota mon
20	76	10.7	571	1 DHM1_METME	P88539 methyllobil
21	73	10.3	776	1 PSAB_HETTR	09x92 heterocapsa
22	70.5	9.9	540	1 PPBI_RAT	P15683 rattus norv
23	70	9.9	207	1 3MGH_LISMO	P58621 listeria mo
24	70	9.9	441	1 GUN2_THEFU	P26222 thermomonas
25	70	9.9	753	1 YBHJ_ECOLI	P57574 escherichia
26	69.5	9.8	400	1 DACC_ECOLI	P08506 escherichia
27	69.5	9.8	661	1 PSAB_PROHO	P88387 prochloroth
28	69	9.7	640	1 INV1_CAPAN	P93761 capsicum an
29	68.5	9.6	417	1 YDHI_ECOLI	P73750 escherichia
30	68.5	9.6	417	1 HEM2_SEIMA	P45623 selaginella
31	68.5	9.6	524	1 P60_LISIV	001837 listeria iv
32	68.5	9.6	864	1 STAZ_PIG	002799 sus scrofa
33	68	9.6	159	1 MP63_MYCTU	P97175 mycobacteri

34	68	9.6	347	1 YA99_MYCPN	P75593 mycoplasma
35	68	9.6	718	1 YOR4_CORNL	P54122 cornebacte
36	68	9.6	740	1 PSAB_ANASP	P58565 anabaena sp
37	68	9.6	1515	1 GLTB_AZOB	005755 azospirillum
38	67.5	9.5	146	1 AZUP_ALCPA	P04377 alcaligenes
39	67.5	9.5	261	1 TPIS_MYCLE	P46711 mycobacteri
40	67.5	9.5	734	1 PSAB_CYARA	09f166 cyanidium c
41	67.5	9.5	851	1 STAZ_HUMAN	P52630 homo sapien
42	67.5	9.5	974	1 PHS2_SOLTU	P53535 solanum tub
43	67	9.4	331	1 DIV_ECOLI	P15286 escherichia
44	67	9.4	485	1 BIND_STREPN	P23118 strongyloce
45	67	9.4	500	1 YFJ1_YEAST	P43601 saccharomyc

ALIGNMENTS

RESULT	ID	STANDARD	PRT	757 AA
1	DHET_GLUOX			
AC	005542			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit I).			
GN	ADHA.			
OS	Gluconobacter oxydans (Gluconobacter suboxydans).			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae; Gluconobacter.			
OX	NCBI_TaxID=442;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.			
RC	STRAIN-IFO 12528;			
RX	MEDLINE=97208225; PubMed=9055427;			
RA	Kondo K., Horinouchi S.;			
RT	*Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus.*			
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).			
CC	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.			
CC	-1- COFACTOR: PO4 AND HEME.			
CC	-1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: D86375; BAA19753.1; -			
DR	HSSP: 0924J7; 1FLG.			
DR	InterPro: IPR001479; Bac_POO.			
DR	InterPro: IPR002372; Bac_POO_repeat.			
DR	InterPro: IPR000345; CytC_heme_bind.			
DR	Pfam: PF01011; Bacterial_POO_6.			
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C_1.			
KW	Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.			
FT	SIGNAL	1	34	
FT	CHAIN	35	757	ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT	MOD_RES	35	35	PYRROLIDONE CARBOXYLIC ACID.
FT	ACT_SITE	342	342	BASE (POTENTIAL).

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 757 AA; 82968 MW; 39B9F90EB3947581 CRC64;

Query Match 29.7%; Score 211; DB 1; Length 757;
 Best Local Similarity 41.2%; Pred. No. 6,1e-13;
 Matches 56; Conservative 20; Mismatches 52; Indels 8; Gaps 5;

QY 7 LMASGALALLAA--PAPVQV---TPVDELIANPPAGEMISYGOENYRHSPLQITTT 61
 Db 16 LLSCAAALAFSAVPAVPAEDPTGTATISSDNGGHP-GDWLSYGRSYSEORISPLDQINT 74
 QY 62 ENVGQLQVWARGMOPKQV-TPLIHGVMTLANPGDVIOAIDAKTGDILMEHRRLP- 119
 Db 75 EAVGKLKLMHYDDLTNNGOESTPLIVGVMTATNMSKMAIDATGKLMSIDPKVPG 134
 QY 120 NIATLNSFEPTRGMA 135
 Db 135 NIADRGCCDTVNRGAA 150

RESULT 2

DHET_ACEPO STANDARD; PRT; 738 AA.

AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RA MEDLINE=91159482; PubMed=2001402;
 RA Tanaka T., Fukaya M., Takemura H., Takemura K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D00635; BAA00528.1; -
 DR PIR: S14270; S14270.
 DR HSSP: 0924J7; 1F1G.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT BINDING 650 650 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 26.3%; Score 187; DB 1; Length 738;
 Best Local Similarity 32.4%; Pred. No. 1.3e-10;
 Matches 44; Conservative 23; Mismatches 59; Indels 10; Gaps 3;

QY 10 SAGLALLAPAPVQVTPVDE-----LANPPAGEMISYGOENYRHSPLQITTT 61
 Db 17 TAGTICALISGATVMSADGGGATGEALITHDDHPGMMYGRYSQGRSPPLDQINR 76
 QY 62 ENVGQLQVWARGMOPKQV-TPLIHGVMTLANPGDVIOAIDAKTGDILMEHRRLP- 119
 Db 77 SNVGNLKLAWYLDLTNNRGOESTPLIVGVMTATNMSKMAIDATGKLMSIDPKVPG 136
 QY 120 NIATLNSFEPTRGMA 135
 Db 137 NIADRGCCDTVNRGAA 152

RESULT 3

DHET_ACEEU STANDARD; PRT; 739 AA.

AC Q44002; 007952;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconacetobacter.
 OX NCBI_TaxID=33995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RA Thurner C.A.K.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X82894; CAA58066.1; -
 DR EMBL: Y09480; CAA70688.1; -
 DR HSSP: 0924J7; 1F1G.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.

us-08-934-506a-5_copy_1_135.rsp

DR
DR HSPD: JS0326; JS0326.
DR PIR: O94J7; 1FIG.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_PQQ_6.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
DR

KW	Oxidoreductase; POO; Heme; periplasmic; Membrane; Signal.
FM	1
KT	35
ET	SIGNAL
ET	CHAIN
ET	ACT SITE
ET	36
ET	343
ET	742
ET	343
ET	ALCOHOL DEHYDROGENASE [ACCEPTOR].
ET	BASE (POTENTIAL).

```

FT BINDING 649 649 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SO SEQUENCE 742 AA; 81521 MW; 96C6C9268DBA8B25A CRC64;

Query Match 25.2%; Score 179; DB 1; Length 742;
Best Local Similarity 36.2%; Pred. No. 7.8e-10;
Matches 46; Conservative 19; Mismatches 56; Indels 4; Gaps 3

QY 11 AGAIIALLAPAFACQVTPYDELL--ANPAGEWISYQNGQENYRHSPLQTITENYGOLQ 68

```

69 LVMARGMQPGKVQ-VTPLIHGVWYLANPQDVIAIDAKTGDLEIWEHRQLP-NIATLNS 122
| : : | ||| : : : | | : : : : : |
82 LI GVMTDMPGCGEATPI VVNGITVATETWSKMFALDAATCKILIMVDYDKVPENIADRGC 147

QY	127	FGEPTRG	133
DB	143	CDTVNRG	149
RESULT	5		
DHG_ECOLI			
ID	DHG_ECOLI	STANDARD;	PRF;
AC	P156727		796 AA.

DT 01-APR-1990 (rel. 14, Created)
 DT 01-NOV-1997 (rel. 35, last sequence update)
 DT 16-OCT-2001 (rel. 40, last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline:quinone] (EC 1.1.99.17).
 GN GCD OR B0124.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Escherichia*.
 OR NCBI:taxid=562.

RN SEQUENCE FROM N.A.
 RP STRAIN-K12;
 RC MEDLINE=91035240; PubMed=2228962;
 RX Cleon-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
 RA "Cloning, mapping, and sequencing of the gene encoding Escherichia
 RT coli quinoprotein glucose dehydrogenase-";
 RI *Bacteroid* 172-6308-6315(1990)

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE=93123180; PubMed=8419307;
RA Yamada M., Asosaka S., Sailer M.H. Jr., Yamada Y.;
RT "Characterization of the qcd gene from *Escherichia coli* K-12 W3110
and regulation of its expression".

RL J. Bacteriol. 175:308-311(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=94261430; PubMed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of
 110,037,192,642 bp, position #

RL Nucleic Acids Res. 22:1637-1639(1994).

[4]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655:
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN
 RP TOPOLOGY.
 RX MEDLINE=93286127; PubMed=8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96128046; PubMed=8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *coli* modelled on that of methanol dehydrogenase from *Methylobacterium*
 RT *extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC
 CC -1- COFACTOR: PQQ.
 CC
 CC -1- SUBUNIT: MONOMER.
 CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: X51323; CA35706.1; -;
 DR EMBL: D12651; BAA02174.1; -;
 DR EMBL: D26562; CAB20298.1; -;
 DR EMBL: AE000122; AAC73235.1; -;
 DR PIR: JVO107; JVO107.
 DR HSSP: P38539; 4AAH.
 DR Ecogene; BG10369; gcd.
 DR InterPro; IPR001479; Bac_PQQ.
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ_7.
 DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase, PQQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 RN
 FT DOMAIN 1 10
 FT TRANSSEM 11 37
 FT DOMAIN 38 40
 FT TRANSSEM 41 58
 FT DOMAIN 59 62
 FT TRANSSEM 63 81
 FT DOMAIN 82 95
 FT TRANSSEM 96 110
 FT DOMAIN 111 118
 FT TRANSSEM 119 141
 FT DOMAIN 142 142
 FT ACT_SITE 466 466
 FT CONFLICT 59 59
 FT CONFLICT 149 156
 FT CONFLICT 193 193
 FT CONFLICT 666 666
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;

Query Match 22.3%; Score 158.5; DB 1; Length 796;
 Best Local Similarity 32.0%; Pred. No. 8,4e-08;
 Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;
 QY 6 LWMAS-----AGALLALPAFAOVTPTDILLANPAGEWISVGNQENRRHSPLQ 58
 DB 135 LTMAGFNDPQELNGTISADATPAEA-ISPVADO-----DMPAYGRQEGORSPSLQ 185
 QY 59 ITTENVGQLQLVAA-----RCMOPGV--QVTPLIHGVMTLANPGVIAIDAKTGD 109
 DB 186 INMDVYNHLEKAWFERGVKQPNDEGETINEVTPIKVGDITLVCVAHQRLFLDAASCK 245
 QY 110 LIMEHRQQLPNITLNSFGPEPT-RCMA 135
 DB 246 EKWHYD---PELKTNESEFQHVTCRGVS 269
 RESULT 6
 EXAA_PSEAE
 ID EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC Q924J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=99041560; PubMed=9826187;
 RA Diehl A., Wintzingerode F., Goerlich H.;
 RT "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
 RA Gardner R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiser J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=9913751; PubMed=10075429;
 RA Schobert M., Goerlich H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerlich H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).

CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AJ009858; CAA08896.1; -
 CC EMBL: AF004624; AAG05370.1; -
 CC EMBL: AF068264; AAC79657.1; -
 CC PDB: 1FUG; 30-AUG-00
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 CC Oxidoreductase; POQ; Periplasmic; Signal; Calcium; 3D-structure;
 KW Complete proteome.
 KW SIGNAL 1 34
 FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SO SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;
 Query Match 19.3%; Score 137; DB 1; Length 623;
 Best Local Similarity 31.3%; Pred. No. 7.9e-06;
 Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;
 QY 3 PTSLMASAGALAL--LAAPAFQVTPVTELLAN--PPAGEWISYGOQENYRHSPLT 57
 DB 9 PAGLRPSLHCLARVALGSGALADYWDIANDDKTGDVLOQGMGTHQMRMPLK 68
 QY 58 QITENNGOLQVWARGN---QPKVQVTPPLIHGVMYLANPGDVIAIDAKTGDLIWE 113
 DB 69 QVNDNVKELTPAMVSYSRSGDEKQGO-ESQATVSDGVIVYASYSRLFALDAKTKRLMT 127
 QY 114 HRROLPN 120
 DB 128 YNHRLLP 134
 RESULT 7
 DHG_GLUOX STANDARD: PRT: 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RT Cleton-Jansen A. M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP STRAIN=LMD 79.41;
 RX GOOSEN N.;

RL submitted (FEb-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 CC PIR: S17716; QPKEX.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 KW SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT SITE 470 470
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 SO SEQUENCE 808 AA; 87567 MW; 0F416DDA78652445 CRC64;
 Query Match 18.9%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 2.1e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;
 QY 13 ALALAAPAFQVTPVTELLA-----NPPAGEWISYGOQENYRHSPLQIQT 61
 DB 128 AVLALFASLFTDPHDISGELPTQIANASPADPDNVASEHAGRTQAGDRMSPLQIQA 187
 QY 62 ENYGOLQVW-----ARCMQPKV--QVTPPLIHGVMYLANPGDVIAIDAKTGDLI 112
 DB 188 TNVSNLKVAMHHTKDMMSNDPGEQTNENATPIEFNNITLYMCSLHQKLFAYDQATGNVAK 247
 QY 113 EHRROL 118
 DB 248 VYDPKL 253
 RESULT 8
 DHGA_ACICA STANDARD: PRT: 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDHA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OC NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=LMD 79.41;
 RX MEDLINE=88289368; PubMed=3399393;

RA Clifton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from *Acinetobacter calcoaceticus*.";
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOCOETICUS CONTAINS TWO DIFFERENT
 CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X07235; CAA30222.1; -
 CC PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KM Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 SQ SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
 Query Match 17.3%; Score 123; DB 1; Length 801;
 Best Local Similarity 30.7%; Pred. No. 0.00024;
 Matches 39; Conservative 16; Mismatches 48; Indels 24; Gaps 5;
 QY 20 PAFQVPTVDELNPAGEMISYGNQENYRHSPLTQTTENVGOLQVWARGQPGK 79
 DB 152 PETAQAVPVAAE-----SDWPATGRTQAGVATYSPKQINDQNVADLAVAT--LRTGD 202
 QY 80 V-----QVPLIHGQVATLPDGVIAIDAKTGDLIWEHRRQIPNATLNSFG 128
 DB 203 LKTDNDGSETTQVTPIKIGNNFICTAHQQLAIDPATGKEKW--RDPKLTDKSPQ 259
 QY 129 EPT-RGM 134
 DB 260 HLTGRCV 266
 RESULT 9
 DHM1_PARDE STANDARD; PRT; 631 AA.
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MEDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RX MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stouthamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RT structural gene from *Paracoccus denitrificans*.";
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 CC TOTAL CELL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M1739; AAA83366.1; -
 CC HSSP: P38539; AAAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KM Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631
 FT DISULFID 135 136
 FT DISULFID 418 447
 FT ACT_SITE 335 335
 SQ SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;
 Query Match 16.8%; Score 119.5; DB 1; Length 631;
 Best Local Similarity 32.18%; Pred. No. 0.00041;
 Matches 45; Conservative 26; Mismatches 54; Indels 15; Gaps 9;
 QY 4 TSLWASAGATALL-AAAFAQVPTDEL--LANPAGEMISYGNQENYRHSPLQIT 60
 DB 12 SSLAAVAVMGLAVLTATATA-----NDQVLELAKDPA-NWVATGRDINAQNTSEMTDIN 65
 QY 61 TENVGOLQVW--ARGKQPGKVQVTPPLIHGQVWLANP-GDVIAIDA-KTGDLIWEHR- 115
 DB 66 KENVKQLRPAMSESTGVILGH-EGPILVVGDRMFTHFPFPTTALDNLNPKILMQNPK 124
 QY 116 RQPLNATLNSFGEPTRGMA 135
 DB 125 KQNPARTVACCDVYVNRGLA 144
 RESULT 10
 DHM1_METOR STANDARD; PRT; 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MEBH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.

DB 67 LRPANFSTGLNGH-EGAPLVGDKMYHTSPNNTFALGLDDPGTILMOKDKONPAA 125
 QY 122 ATLNSEPTRGMA 135
 DB 126 RAVACCDLVNRGLA 139

RESULT 12
 ID XOXF_PARDE STANDARD; PRT; 600 AA.
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harms N.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE OF 380-600 FROM N.A.
 RC STRAIN=PD 1235;
 RX MEDLINE=92041583; PubMed=1657873;
 RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
 RA Stouthamer A.H.;
 RT Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c551 of Paracoccus denitrificans and characterization of
 RT the mutant strain."
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1- COFACTOR: POQ (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U34346; AAC44555.1;
 DR EMBL: M75583; AAA25574.1;
 DR PIR: A41378; A41378.
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ; 7.
 KW Oxidoreductase; POQ; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 600
 FT ACT_SITE 318 318
 FT SEQUENCE 600 AA; 65159 MW; DCA96F1BC5A3CE CRC64;
 FT PUTATIVE DEHYDROGENASE XOXF.
 FT BASE (POTENTIAL).
 FT ACT_SITE 318 318
 FT SEQUENCE 600 AA; 65159 MW; DCA96F1BC5A3CE CRC64;

Query Match 13.9%; Score 99; DB 1; Length 600;
 Best Local Similarity 27.5%; Pred. No. 0.038;
 Matches 38; Conservative 25; Mismatches 53; Indels 22; Gaps 7;

QY 10 SAGALALLAFAFOVPTDELALNP-----PAGEWISTGONCEYRHSPLQITTE 62
 DB 6 NAACIALINSGTAA-----LANEORAGROAPQWALQMGDYANTRYSTLQIKNMD 56
 QY 63 NVGOLQIVM-ARGMOPGVQVYPLIHGCVYLANP-GDVIOAIDAK-TGDLWEHR-RQ 117
 DB 57 NVKDLRVAMTSTGYLRGH-BESPLVIGDVYVHTPPRVFALDLNNGKILMYEPOQ 115
 QY 118 LPNIATLNSEPTRGMA 135
 DB 116 DENVIAMVMSCTDYVRLGS 133

RESULT 13
 ID QUITA_ACICA STANDARD; PRT; 809 AA.
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=BD413 / ADP1;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsomore D.A., Ornstion L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quita, the structural gene for quinate-shikimate
 RT dehydrogenase."
 RL J. Bacteriol. 176:7659-7666(1994).
 RP [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN=BD413 / ADP1.
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsomore D.A., Ornstion L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus."
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: quinate + pyrroloquinoline-quinone = 5-
 CC -1- CATALYTIC ACTIVITY: shikimate + pyrroloquinoline-quinone = 3-
 CC -1- CATALYTIC ACTIVITY: shikimate + pyrroloquinoline-quinone = 3-
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY, FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PYROCATECHONATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PYROCATECHONATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L05770; AAC37161.1;
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ; 7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT TRANSMEM 61 88
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 FT SEQUENCE 809 AA; 88196 MW; 71F67CBEA62BFCB CRC64;

Query Match 13.9%; Score 99; DB 1; Length 809;
 Best Local Similarity 29.0%; Pred. No. 0.053;
 Matches 42; Conservative 15; Mismatches 54; Indels 34; Gaps 6;

QY 3 PISLWASAGALALIAA-PAPAQVPTVEDELLANPAGEMISYGNONEYRHSPLTQIT 61
 DB 148 PHEVTAASEEELPLVVDPAKKOV-----NMWHDYDAGSGSRVALDQINR 193
 QY 62 ENYGOLQIYV-----ARGMOPG-KVQVYPLIHGVMYLANPGDVYIQADIKATGLIME 113
 DB 194 NNSVSKIEKAMRPTGDFGTGTGAGADQMTPLQYGNKVFELCPHNNIFADIDSGKOLMK 253
 QY 114 HRRQLPNIATLNSFG---EPTRGMA 135
 DB 254 -----AEVNSTADAMERCGRVA 270
 RESULT 14
 QUTAL_XANCU STANDARD; PRT; 790 AA.
 ID QUTAL_XANCU
 AC 09XD78;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable quinolate dehydrogenase [Pyrrroloquinoline-quinone]
 DE (EC 1.1.99.25).
 GN QUMA.
 OS Xanthomonas campestris (pv. juglandis).
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 NCBI_TaxID=44291;
 RX MEDLINE=20063481; PubMed=10594704;
 RA Lee Y.-A., Lo Y.-C., Yu P.-P.;
 RT "A gene involved in quinolate metabolism is specific to one DNA homology
 RT group of Xanthomonas campestris."
 RL J. Appl. Microbiol. 87:649-658(1999).
 CC -1- CATALYTIC ACTIVITY: quinolate + pyrroloquinoline-quinone = 5-
 CC dehydroquinolate + reduced pyrroloquinoline-quinone.
 CC -1- COFACTOR: PQO (BY SIMILARITY).
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP, THIS PATHWAY
 CC ALLIOMS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF109471; AAD38453.1; -;
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_6.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; FALSE_NEG.
 KW Oxidoreductase; PQO; Quinolate metabolism; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 77 94 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 SQ SEQUENCE 790 AA; 82896 MW; B75F29B52A49FE6F CRC64;

Query Match 12.9%; Score 91.5; DB 1; Length 790;
 Best Local Similarity 27.6%; Pred. No. 0.28;
 Matches 40; Conservative 20; Mismatches 60; Indels 25; Gaps 7;
 OY 9 ASAGALALLA-AFAFAQVPTVDE-----LLANPPAG---EWISYGNONEYRHSPLTQI 59

DB 110 ATVALAVAGIGMFPHPVPAAGNAGPMTAVPGSVQNWSAVYNTDGGSRFAALDOI 169
 QY 60 TTEVNGOLQIYVARGMOPGKV-----QVPLIHGVMYLANPGDVYIQADIKATGD 109
 DB 170 NRSN-GRPAAGSPGPTTGTGTGANSNGAEQQLPPLQYGEVFLCTHNNILALDASTGK 228
 QY 110 LIWEHRRQLPNIATLNSFGEPTRG 134
 DB 229 QLM--RRE---INATSSVWQRCRGL 248
 RESULT 15
 CU08_LOCM1 STANDARD; PRT; 148 AA.
 ID CU08_LOCM1
 AC P11734;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cuticle protein 8 (LM-8) (LM-ACP 8).
 DE Locusta migratoria (Migratory locust).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthoptera; Orthopteroidea; Caelifera;
 OC Acridoidea; Acrididae; Acrididae; Locusta.
 GN Acridomorpha; Acrididae; Acrididae; Locusta.
 NCBI_TaxID=7004;
 RX MEDLINE=90073593; PubMed=2590176;
 RA Klarskov K., Hoelrup P., Andersen S.O., Roepstorff P.;
 RT "Plasma-desorption mass spectrometry as an aid in protein sequence
 RT determination. Application of the method on a cuticular protein from
 RT the migratory locust (Locusta migratoria).";
 RL Biochem. J. 262:923-930(1989).
 CC [2]
 CC SEQUENCE OE 1-56.
 CC MEDLINE=86108304; PubMed=3943519;
 RA Hoelrup P., Andersen S.O., Roepstorff P.;
 RT "Isolation, characterization, and N-terminal sequence studies of
 RT cuticular proteins from the migratory locust, Locusta migratoria.";
 RL Eur. J. Biochem. 154:153-159(1986).
 CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
 CC CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
 CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-LAV) REPEATS FOUND THROUGHOUT THE
 CC PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
 CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
 CC -1- SIMILARITY: CONTAINS 1 CUTICLE CONSENSUS DOMAIN.
 CC PIR: B24802; B24802.
 DR PIR: S05638; S05638.
 DR InterPro: IPR000618; Insect_cuticle.
 DR Pfam: PF00379; Insect_cuticle; 1.
 DR PRINTS: PR00947; CUTICLE.
 DR PROSITE: PS00233; CUTICLE; 1.
 KW Structural protein; Cuticle; Repeat.
 FT REPEAT 16 19
 FT REPEAT 22 22
 FT REPEAT 28 31
 FT REPEAT 37 40
 FT REPEAT 44 47
 SQ SEQUENCE 148 AA; 15224 MW; C6EAD2C7C593AC66 CRC64;

Query Match 11.1%; Score 78.5; DB 1; Length 148;
 Best Local Similarity 23.1%; Pred. No. 0.77;
 Matches 24; Conservative 17; Mismatches 20; Indels 43; Gaps 3;
 OY 8 WASAGALALLAPAPAQVPTVEDELLANPAGEMISYGNONEYRHSPLTQITTEVNGOL 67
 DB 27 YAAVPAKAVVAPAAVAPAAVAPAAVAPAAVAPAAVAPAAVAPAAVAPAAVAPAAVAPAA 66
 QY 68 QLVWARGMOPGKVQVYPLIHGVMYLANPGDVYIQADIKATGLI 111
 DB 67 -----VHD-----AHTGDIKQOSEARDGIVV 87

Fri May 24 11:27:20 2002

us-08-934-506a-5_copy_1_135.rsp

Page 10

Search completed: May 24, 2002, 10:15:50
Job time: 768 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:31 ; Search time 93.93 seconds

(without alignments)
248.635 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Sequence: 1 MKPTSLMAGALALALAP.....RQPNITATINSFGEPTRGMA 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	268	37.7	608	2	Q93RE9 pseudogluc
2	191	26.9	742	2	O53362 acetobacter
3	178.5	25.1	691	2	O9AP95 pseudomonas
4	174	24.5	623	2	O9AGW3 pseudomonas
5	168.5	23.7	698	2	O9KH03 alcaligenes
6	167	23.5	786	2	P95466 pantoea cit
7	165.5	23.3	708	2	O46444 comamonas t
8	156	22.0	695	2	O9F9U2 pseudomonas
9	155	21.8	785	16	O98KF6 rhizobium m
10	146.5	20.6	777	16	O92RB3 rhizobium m
11	143	20.1	803	16	O91115 pseudomonas
12	140.5	19.8	601	2	O9EYW8 rhizobium m
13	140.5	19.8	601	16	O92WY9 pantoea cit
14	121	17.0	790	2	O9X2S5 methylobact
15	119	16.8	601	2	P71509 pseudomonas
16	118.5	16.7	695	2	O934G0 pseudomonas

17	111.5	15.7	629	2	O9A048 methylobact
18	105.5	14.9	633	2	O24759 hyphomicrob
19	93.5	13.2	639	2	P77931 pseudomonas
20	93.5	13.2	644	2	O52551 pseudomonas
21	89	12.5	499	2	O51842 thiodacillu
22	86	12.1	613	17	O28873 archaeoglob
23	84.5	11.9	599	2	O91935 methylovoru
24	83.5	11.8	2832	2	O93N05 agrobacteri
25	82.5	11.6	485	17	O96XP7 sulfolobus
26	82	11.5	668	16	O9RRH3 delnococcus
27	81	11.4	680	5	O9N9G7 phallusia m
28	80.5	11.3	647	16	O9HUB1 pseudomonas
29	79.5	11.2	846	2	O86092 methylovoru
30	78.5	11.1	509	16	P96368 mycobacteri
31	78	11.0	207	16	O92D89 listeria in
32	76	10.7	306	16	O9RY72 delnococcus
33	76	10.7	386	16	O9KTW8 vibrio chol
34	76	10.7	573	2	O59540 methylophil
35	76	10.7	4809	2	O93HH0 streptomyce
36	75.5	10.6	407	16	O9PG38 xylella fas
37	75.5	10.6	684	2	O9S282 streptomyce
38	75.5	10.6	1419	5	O9VYL3 P73139 synecocyst
39	75.5	10.6	3972	16	P73139 synecocyst
40	75	10.6	183	16	O9CKL1 pasteurilla
41	74.5	10.5	969	16	O9RZ15 delnococcus
42	74	10.4	294	2	O50152 streptomyce
43	74	10.4	380	16	O9HXU7 pseudomonas
44	74	10.4	525	16	O931E9 stapyllococ
45	74	10.4	755	2	O9WWAL agrobacteri

ALIGNMENTS

RESULT 1
ID Q93RE9 PRELIMINARY: PRT: 608 AA.
AC Q93RE9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharoketogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_Taxid=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shibata T., Saito Y.;
RT "Alcohol dehydrogenase."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1;
SQ SEQUENCE 608 AA: 65101 MW: 0ACBC97AE11BA570 CRC64:

Query Match 37.7%; Score 268; DB 2; Length 608;
Best Local Similarity 40.4%; Pred. No. 7.3e-18;
Matches 59; Conservative 21; Mismatches 46; Indels 20; Gaps 2;
QY 10 SAGALALAAPAPAO-----VTPVDELLANPPAGEWISYGOEN 50
DB 16 STALIALSGPAPAFQHANAAPERSKAGSAIENFQVTFADDLAKRPANWPIIRGYOG 75
QY 51 YRHSPLQITTEENVGQLQLYWARGMGKQVPLIHIDGYMYLANPQVQALDAKTGDL 110
DB 76 WGYSPLDQINKNDVGDQLVWSRTMEPSNEGAIAIVNGVIFLGNVDVLOALDKTGSL 135
QY 111 IWEHROLPNLAT-LNSFGEPTRGMA 135
DB 136 IWEHRLKLPASKRINSLGAKRSIA 161

```

RESULT 2
Q53362 PRELIMINARY; PRT; 742 AA.
ID 053362
AC 053362: 044159;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DE 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus (Acetobacter turbidans).
OC Bacteria: Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC11380;
RA MEDLINE=94042848; PubMed=8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus."
RL J. Bacteriol. 175:6857-6866(1993).
DR EMBL; D13893; BAA40252.1; -.
DR HSSP; Q924J7; 1FLG.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_6.
DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match
Best Local Similarity 26.9%; Score 191; DB 2; Length 742;
Matches 47; Conservative 18; Mismatches 58; Indels 4; Gaps 3;

QY 11 AGALALLAPAFQVPTVDLDEL-ANPPAGEWISYSGONENYRHSPLQITTEENGOLO 68
DB 23 AALPYAAVADPARADGGNGTGEAIHDDHPENMLSTGRYSRSPLOQINRSVNGDLK 82
QY 69 LVWARGMOGKVO-VTPLIHGCVMTLANPGDVIQAIIDAKTGDLIWEHRRLP-NITATLS 126
DB 83 LAWYTLDTNRGEGATPLVVDGIMVATNWSKMEALDAATGKLIMQYDKVPGNIADKGC 142
QY 127 FGEPTFG 133
DB 143 COTVNGC 149

RESULT 3
Q9AF95 PRELIMINARY; PRT; 691 AA.
ID 09AF95
AC 09AF95:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BDH.
OS Pseudomonas butanovora.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF355798; AAK27220.2; -.
SQ SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match
Best Local Similarity 25.1%; Score 178.5; DB 2; Length 691;
Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

```

```

QY 39 GEMISYSGONENYRHSPLQITTEENYGOLOLVWARGMOPGK-VQYTPLIHGCVMTLANPG 97
DB 41 GEMRTGYDAGTRYSPLQITPDNAKELGLVMSYDLESSRGVATPTIVDGVMTVTAAPW 100
QY 98 DVIQATDAKTGDLIWEHRRLP 119
DB 101 SVVHALDVSGKRLMTYDPEVP 122

RESULT 4
Q9AGW3 PRELIMINARY; PRT; 623 AA.
ID 09AGW3
AC 09AGW3:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1; -.
DR HSSP; Q924J7; 1FLG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_4.
KW Signal.
SQ SIGNAL 1
SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match
Best Local Similarity 24.5%; Score 174; DB 2; Length 623;
Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLWASAGALLAPAFQVPTVDLDEL-PPAGEWISYSGONENYRHSPLQI 59
DB 7 KPFALRAIVATAAALSLPA-AAVTDVTMEDIANDKHTGDVLTGGLAAGHSPLKAI 65
QY 60 TTEWVGQLQLYMA-----RGMPQKQVQYTPLIHGCVMTLANPGDVIQAIIDAKTGDLI 111
DB 66 NTDVNAVNLVPAWSFSFGEGKRGQEE--AOV--LVHDGVYATASYSRIFAIDARSGKRL 120
QY 112 WEHRRLP 120
DB 121 WEYNAKLP 129

RESULT 5
Q9KH03 PRELIMINARY; PRT; 698 AA.
ID 09KH03
AC 09KH03:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TERAHYDROFURFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria: Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21125557; PubMed=11222593;
RA Zarni G., Schrader T., Andreesen J.R.;
RT "Catalytic and Molecular Properties of the Quinolomoprotein
RT tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha

```


CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POO, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CA57464.1; -.
 DR HSSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_C; 1.
 DR PRINTS: PR00605; CYTOCHROME_C.
 KW Signal; POO; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635 QUTINHAEMOPROTEIN ETHANOL DEHYDROGENASE
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACCB3 CRC64;

Query Match 23.3%; Score 165.5; DB 2; Length 708;
 Best Local Similarity 31.8%; Pred. No. 9.3e-08;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

OY 3 PTLISMSAGALLAAPAFQVTPVY-----DELANPP-AGEWISYGOQEN 50
 DB 12 PGRWMLAACLG--SAAAFQGTGPAQAQAAAVGVVDGDFIRMAARTPMPRTIGVDYAE 69
 OY 51 YRHSPLQITTEYNGOLVWARGMPGK-VQVPLIHGVMYLANGDVIOAIDAKTGD 109
 DB 70 TRYSLRDLQINMANVKDGLMSVYLESTREVEATPVVDGIMYASMSVYHAIIDRTGN 129
 OY 110 LIWEHRQL 118
 DB 130 RIWTYDPOI 138
 RESULT 8
 O9F9U2 PRELIMINARY; PRT; 695 AA.
 AC Q9F9U2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainerd J.B.,
 RA Terwilliger T.C.;
 RT Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2;
 RT Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; -.
 DR HSSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR003088; Cyt_C1.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_C; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 SO SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 22.0%; Score 156; DB 2; Length 695;
 Best Local Similarity 33.6%; Pred. No. 7.7e-07;
 Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;

OY 15 ALAAPAFQVTPVDEL--ANPAGWISYGOQENYRHSPLQITTEYNGOLVWA 72
 DB 16 ALLVA-AGAAVAYDEAIAISQDSSEMLSHQRTAEDRFSPKOIDAGNKGILAWY 74
 OY 73 RGMQPK-VQVPLIHGVMYLANPGDVIOAIDAKTGDILWEHRQL 118
 DB 75 LDLENNRGLFATPLVSDGYLVAISLMSRVAVDLRSGRKLMQFDPPQY 121

RESULT 9
 O98KF6 PRELIMINARY; PRT; 785 AA.
 AC Q98KF6;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN MLL1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RC MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RT DNA Res. 7:331-338(2000).
 DR EMBL: AP002997; BAB48856.1; -.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO; 7.
 KW Complete proteome.
 SO SEQUENCE 785 AA; 84083 MW; IDC224B7A1C25470 CRC64;

Query Match 21.8%; Score 155; DB 16; Length 785;
 Best Local Similarity 31.7%; Pred. No. 1.1e-06;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;

OY 8 WASAG-----ALALAAPAFQVTPYTD-----ELANPAGE 40
 DB 111 WARRGLAPDGPAPLILAVLSLAVAGYSMTADPKGIDGALDTOKVIPKANLGVDPAGE 170
 OY 41 WISYGOQENYRHSPLQITTEYNGOLVWARGMPGK-----QVPLIHG 89
 DB 171 WHYGRTOFGORYSPLDITPDNVANLQPAWT--YRIGDVAGPDDIGETTYQVPLIAGD 228
 OY 90 WMYLANPGDVIOAIDAKTGDILWEH 114
 DB 229 TLYICTPHNFAIVDAATGKREKWR 253

RESULT 10
 O92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYROLOQUINOLINE-OUTINONE) PROTEIN
 DE (EC 1.1.99.17).

Query Match	20.68;	Score 146.5;	DB 16;	Length 777;
Best Local Similarity	38.5%;	Pred. No. 7.5e-06;		
Matches 35; Conservative	9;	Mismatches 34;	Indels 13;	Gaps 2

RESULT	11
Q9115	
ID	Q9115
PRELIMINARY;	
PRT;	803 AA

```

DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-OCT-2001 (TREMBLrel. 16, Last annotation update)
DE      GLUCOSE DEHYDROGENASE.
DE      GCD OR PA2290.
OS      Pseudomonas aeruginosa.
OC      Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX      Pseudomonas.
OX      NCBI_TaxId=287;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 15692 / PAOI;
RX      MEDLINE=20437373; PubMed=10984043;
RA      Stover C.K., Pham X.-Q.T., Errin A.L., Mizoguchi S.D., Warrenner P.,
RA      Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA      Garner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA      Brody L.L., Coulter S.N., Folger K.R., Kas A., Landig K., Lim R.M.,
RA      Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA      Reizer J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT      "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT      opportunistic pathogen.";
RL      Nature 406:959-964(2000).
DR      EMBL: AE004654; AAC05678.1; -.
DR      InterPro: IPR002372; Bac_P00 repeat.
DR      Pfam: PF01011; Bacterial_P00; 7.
KW      Complete proteome.
SQ      SEQUENCE 803 AA; 86217 MW; 7DDF46B0EB5ED249 CRC64;

```

Query Match	20.1%;	Score 143;	DB 16;	Length 803;
Best Local Similarity	26.9%;	Pred. No. 1.7e-05;		

RP SEQUENCE FROM N.A.
RC STRAIN-PM1021.
RA Fenner B.J., Tiwari R.P., Dilworth M.J.;
RT "Regulation of Cl assimilation in *Sinorhizobium meliloti*."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF309488; AAC31643.1; -
DR HSSP: P38539; 4AAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
SQ SEQUENCE 601 AA: 65686 MW: 09F743726B2E2A83 CRC64;

Query Match	19.8%	Score 140.5;	DB 2;	Length 601;
Best Local Similarity	31.8%	Pred. No. 2.1e-05;		
Matches 41; Conservative	29;	Mismatches 50;	Indels 9;	Gaps 7;

[illegible]

RESULT	13	
Q92WY9		
ID	Q92WY9	PRELIMINARY;
		PRT;
		601 AA

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC
 1.1.99.8).
 GN SME20173.
 GN Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Fihon T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhiesler F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA "The complete sequence of the 1,683-kb PSYMB megaplasmid from the N2-
 RT fixing endosymbiont *Shorhizobium meliloti*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AF603642; CAC48573.1;
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424FEED15ADBE CRC64;

Query Match 19.8%; Score 140.5; DB 16; Length 601;
 Best Local Similarity 31.8%; Pred. No. 2.1e-05;
 Matches 41; Conservative 29; Mismatches 50; Indels 9; Gaps 7;
 QY 14 LALIAAPFAPOVPTDEL--LANPAGEWISYGOENYRHSPLQITENYGOLOLVW 71
 DB 8 LAISIGGAQVAFANDELQKLIIDP--NQMAIQGVANLRYSKLDQINKDNVKGLOVAV 66
 QY 72 --ARGMPGKVQVTPPLIHGVMILANP-GVYQAIID-AKTGDLIWEHR-RQLPNIATLNS 126
 DB 67 TTFSTGVLRGH-EGSPLVIGDLMYVHPFNTVYALDLSKDGQIWMKYEKODPNVIPVC 125
 QY 127 FGEPTRGMA 135
 DB 126 CDTVNRGVA 134

RESULT 14
 QYX255 PRELIMINARY; PRT; 790 AA.
 AC Q9X255;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE GUDCOSE DEHYDROGENASE.
 GN GUDH.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.
 OX NCBI_TaxID=53336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1056R;
 RX MEDLINE=99303331; PubMed=10376838;
 RA Pujol C.J., Kado C.I.;
 RT "gdhB, a gene encoding a second quinoprotein glucose dehydrogenase in
 RL *Pantoea citrea*, is required for pink disease of pineapple."
 RL Microbiology 145:1217-1226(1999).
 DR EMBL: AF050503; AAD23735.1;
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00.
 DR Pfam: PF01011; Bacterial_P00.
 DR PROSITE: PS00364; BACTERIAL_P00.2.1.
 SQ SEQUENCE 790 AA; 85681 MW; 88BB7625E0A5F089 CRC64;

Query Match 17.0%; Score 121; DB 2; Length 790;
 Best Local Similarity 32.6%; Pred. No. 0.0024;
 Matches 31; Conservative 13; Mismatches 39; Indels 12; Gaps 3;
 QY 27 PYTDELANPAGEWISYGOENYRHSPLQITENYGOLOLVW-----RGMPGK 79
 DB 152 PATQPSVSAK--DWTAYGTOAGDRYSSLQINSNKNQNVAMTYGSGVKKRPDDGE 208
 QY 80 V--QVTPPLIHGVMILANP-GVYQAIIDAKTGLIIV 112
 DB 209 ITNESTPLKVGMLTYCTAHQILVAIDAATGKEKW 243

RESULT 15
 ID P71509 PRELIMINARY; PRT; 601 AA.
 AC P71509;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MXAP.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=97312011; PubMed=9168622;
 RA Chistoserdova L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RL methylotrophy gene clusters in *Methylobacterium extorquens* AM1."
 RL Microbiology 143:1729-1736(1997).
 DR EMBL: U72662; AAB58890.1;
 DR HSSP: P38539; AAH.
 DR InterPro: IPR002372; Bac_P00.repeat.
 DR Pfam: PF01011; Bacterial_P00.7.
 SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CBB239 CRC64;

Query Match 16.8%; Score 119; DB 2; Length 601;
 Best Local Similarity 29.3%; Pred. No. 0.0027;
 Matches 41; Conservative 29; Mismatches 58; Indels 12; Gaps 8;
 QY 1 MKPTSLMASGALALIAAPFAOVTPTDELANPAGEWISYGOENYRHSPLQIT 60
 DB 1 MRAVHLALGAGLA--ASPALANESVLKG--VANP--AEQVLOTVDYANTRYSKLDQIN 54
 QY 61 TENVGLOLVW--ARGMPGKVQVTPPLIHGVMILANP-GVYQAIIDAKNG-DLIWEHR- 115
 DB 55 ASNVKNLQVAMTFSTGVLRGH-EGSPLVGNIMTVHPFNTVYALDLDGAKIWMKYE 113
 QY 116 RQLPNIATLNSFGEPTRGMA 135
 DB 114 KODPSVIPVMCCDTVNRGLA 133

Search completed: May 24, 2002, 10:17:32
 Job time: 745 sec

Fri May 24 11:27:20 2002

us-08-934-506a-5_copy_1_135.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:02 ; Search time 117.28 Seconds
(without alignments)
357.050 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556
Perfect score: 2032
Sequence: 1 IVANGIVAVAGSTQGYSPFG.....GMQYVALAGGVSYSGSLNS 377

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	579	19	AAW37873 Alcohol and/or ald
2	1794	88.3	579	19	AAW37876 Alcohol and/or ald
3	1777	87.5	579	19	AAW37874 Alcohol and/or ald
4	1736	85.4	578	19	AAW37875 Alcohol and/or ald
5	338	16.6	742	11	AAW05235 Amino acid sequenc
6	332.5	16.4	738	13	AAW20192 ADH complex protei
7	331.5	16.3	738	11	AAW31993 A.aliceteligenses me
8	326.5	16.1	754	21	AAW35987 Sorbitol dehydroge
9	202	9.9	740	20	AAW95019 Sorbitol dehydroge
10	108.5	5.3	530	22	AAU12236 Human poly
11	107	5.3	269	11	AAW07006 Protein G variant.

12	107	5.3	269	15	AAW53295 IgG-binding Strept
13	106.5	5.2	593	11	AAW07014 Protein G variant
14	106.5	5.2	594	12	AAW10005 Streptococcus GX78
15	106	5.2	269	10	AAW94788 Protein G variant.
16	106	5.2	715	22	AAW81738 S. epidermidis ope
17	105.5	5.2	429	22	AAW66542 Human interferon-a
18	105.5	5.2	429	22	AAW93799 Human protein can
19	105.5	5.2	469	21	AAW56578 Human protease sequ
20	105.5	5.2	278	8	AAW70468 Sequence of polype
21	105	5.1	549	21	AAW96642 S. epidermidis ope
22	103	5.0	790	22	AAW67230 Amino acid sequenc
23	102	4.9	391	21	AAW81628 Streptococcus pneu
24	99.5	4.9	770	22	AAW30824 S. epidermidis ope
25	99.5	4.9	910	22	AAW33007 S. epidermidis ope
26	99.5	4.9	696	22	AAW82914 Human protein sequ
27	99	4.9	886	22	AAW93996 Expression vector
28	99	4.8	228	21	AAW10433 Streptococcus Prot
29	98.5	4.8	593	15	AAW62944 C. trachomatis CT8
30	98.5	4.8	1016	22	AAW38901 Drosophila melanog
31	98.5	4.8	7107	22	AAW58144 Chlamydia trachoma
32	98	4.8	524	20	AAW37238 S. aureus Strd pro
33	97.5	4.8	1315	20	AAW08642 Staphylococcus aur
34	97.5	4.8	1349	22	AAW34402 Staphylococcus aur
35	97.5	4.8	1349	22	AAW37584 Fall Armyworm Bt t
36	97.5	4.8	1734	22	AAW20500 Cosmid cHR15 enco
37	97.5	4.8	1673	21	AAW95707 Neisseria meningit
38	97	4.8	2599	21	AAW75098 Novel human diagno
39	97	4.8	366	22	AAW82859 Human tissue trans
40	96.5	4.7	548	21	AAW44918 S. epidermidis ope
41	96.5	4.7	470	22	AAW82701 Human tissue trans
42	96	4.7	470	22	AAW84013 C. trachomatis I2
43	95.5	4.7	505	20	AAW16739 Human GTP-binding
44	95	4.7	505	20	AAW16739 Human GTP-binding
45	95	4.7	547	22	AAW68556

ALIGNMENTS

RESULT 1	AAW37873	standard; Protein: 579 AA.
ID	AAW37873	
XX	AAW37873:	
AC	10-AUG-1998 (first entry)	
XX		
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX		
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulonic acid; L-ascorbic; inhibition.	
XX		
OS	Glucanobacter oxydans.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..23
FT		/note="signal peptide"
FT	Protein	24..579
FT		/note="mature protein"
XX		
XX	EP832974-A2.	
XX		
PD	01-APR-1998.	
XX		
PF	11-SEP-1997;	97EP-0115801.
XX		
PR	19-SEP-1996;	96EP-0115001.
XX		
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
PI		
XX		

DR WPI; 1998-195228/18.
 DR N-PSDB; AAV29051.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1; Pages 35-37; 59pp; English.
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX Sequence 579 AA:
 SQ

Query Match 100.0%; Score 2032; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 5.5e-178;
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVANGYIVAGSTCOYSPFCFVSGHDSATGEELMRNYFIPRAGEEDDETNGDYEARMMT 60
 DB 180 IVANGYIVAGSTCOYSPFCFVSGHDSATGEELMRNYFIPRAGEEDDETNGDYEARMMT 239
 QY 61 GAMQITTYDPTNLVHYGSTAVGPASETORGTGCTLGTNTNFAVRPDTGETIWRHQTJ 120
 DB 240 GWGQITTYDPTNLVHYGSTAVGPASETORGTGCTLGTNTNFAVRPDTGETIWRHQTJ 299
 QY 121 PRDNMOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAAE 180
 DB 300 PRDNMOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAAE 359
 QY 181 TGEFLWARDNTYONMIESIDENGIYTVNEDAILKEIDVEYDVCPTFLGRDPSAALNPD 240
 DB 360 TGEFLWARDNTYONMIESIDENGIYTVNEDAILKEIDVEYDVCPTFLGRDPSAALNPD 419
 QY 241 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPPEKDMIGRIDAIDISTGRTLSV 300
 DB 420 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPPEKDMIGRIDAIDISTGRTLSV 479
 QY 301 ERAAANYSPVLTSGGVLFNSTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGMQ 360
 DB 480 ERAAANYSPVLTSGGVLFNSTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGMQ 539
 QY 361 YVIAAGGVSYSGSLNS 377
 DB 540 YVIAAGGVSYSGSLNS 556

RESULT 2
 AAW37876
 ID AAW37876 standard; Protein: 579 AA.
 XX AAW37876;
 AC
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..23

FT /note= "signal peptide"
 FT 24..579
 FT "note= "mature protein"

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB; AAV29054.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 Claim 1; Pages 44-46; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans
 alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
 or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX Sequence 579 AA:
 SQ

Query Match 88.3%; Score 1794; DB 19; Length 579;
 Best Local Similarity 85.7%; Pred. No. 4.2e-156;
 Matches 323; Conservative 27; Mismatches 27; Indels 0; Gaps 0;

QY 1 IVANGYIVAGSTCOYSPFCFVSGHDSATGEELMRNYFIPRAGEEDDETNGDYEARMMT 60
 DB 181 IVANGYIVAGSTCOYSPFCFVSGHDSATGEELMRNYFIPRAGEEDDETNGDYEARMMT 240
 QY 61 GAMQITTYDPTNLVHYGSTAVGPASETORGTGCTLGTNTNFAVRPDTGETIWRHQTJ 120
 DB 241 GWGQITTYDPTNLVHYGSTAVGPASETORGTGCTLGTNTNFAVRPDTGETIWRHQTJ 300
 QY 121 PRDNMOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAAE 180
 DB 301 PRDNMOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAAE 360
 QY 181 TGEFLWARDNTYONMIESIDENGIYTVNEDAILKEIDVEYDVCPTFLGRDPSAALNPD 240
 DB 361 TGEFLWARDNTYONMIESIDENGIYTVNEDAILKEIDVEYDVCPTFLGRDPSAALNPD 420
 QY 241 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPPEKDMIGRIDAIDISTGRTLSV 300
 DB 421 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPPEKDMIGRIDAIDISTGRTLSV 480
 QY 301 ERAAANYSPVLTSGGVLFNSTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGMQ 360
 DB 481 ERAAANYSPVLTSGGVLFNSTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGMQ 540
 QY 361 YVIAAGGVSYSGSLNS 377
 DB 541 YVIAAGGVSYSGSLNS 557

RESULT 3
 AAW37874

```

ID AAM37874 standard; Protein; 579 AA.
XX
XX AAM37874;
XX
XX 10-AUG-1998 (first entry)
XX
XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
DE
XX
XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
XX aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
XX 2-keto-L-gulononic acid; L-ascorbic; inhibition.
XX
XX Gluconobacter oxydans.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note= "signal peptide"
FT 24..579
FT Protein /note= "mature protein"
XX
XX EP832974-A2.
XX
XX 01-APR-1998.
XX
XX 11-SEP-1997; 97EP-0115801.
XX
XX 19-SEP-1996; 96EP-0115001.
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX WPI; 1998-195228/18.
XX N-PSDB; AAV29052.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
XX dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 38-40; 59pp; English.
XX
XX This is the amino acid sequence for the Gluconobacter oxydans
XX alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
XX or recombinant organisms can be used to convert suitable substrates
XX to aldehydes, ketones or carboxylic acids, especially to convert
XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
XX converted to L-ascorbic acid by standard procedures. The derivatives
XX of ADH enzymes have desired substrate specificity, higher affinity
XX to a substrate, lower affinity to an inhibitory compound, higher
XX stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 579 AA:
SQ

```

```

Query Match 87.5%; Score 1777; DB 19; Length 579;
Best Local Similarity 85.3%; Pred. No. 1.5e-154;
Matches 318; Conservative 29; Mismatches 26; Indels 0; Gaps 0;

```

```

OY 1 IVANGVIVAGSTCOYSPFCFVSGHDSATGEELMRYFTIPRAGEEDENGYEARMMT 60
DB 180 ivangviavagstcypfcfvgshdsatgeelwryfiprgeedetkwyndyeyrmt 239
OY 61 GAWGQITTYDPTNMLVHYGSTAVGASSETORGPFGTLTGNTFRFAVRPDTGLWRHQT 120
DB 240 gawgqittdpntmlvhygstavgasetorgpfgtlgtntfrfavrpdgtglwrhqt 299
OY 121 PRONMDECFEFMAVTVNVQVSTMEGLOSTINRNATGERRVLTGVPCKTGMWQFDE 180
DB 300 prdnmdecfefmavtvnvqvstmeeglostinrnatgerrvltgvpcktgmwqfdae 359
OY 181 TGEFIARDTNYONMIESIDENGIIVTNEDAILKEIDVEYDVCPTFAGSRDWPSSALNDP 240
DB 360 tgefiaradtnyoniiesidengiiivtnedailkeidveydvcptfagsrdwpsalndp 419

```

```

OY 241 SGVIFPLNNVCDMAVDQEFISMDYNTSNVTKLPKDKMIGRIDAIDISTGRTIMSV 300
DB 420 tgyifplnnvcdmavdqefismdyntsntklpdkdkmigridaidistgrtimsv 479
OY 301 ERANANSPVLTSTGCVLPFGNGDTRFRALSOEFGELTNTRLATVAGSAISEYDVGQ 360
DB 480 ereananspvlstgcvlpfngdtrfralsoefgelntntrlatvagsaaiseyldvgq 539
OY 361 YVATAGGVSYS 373
DB 540 yvataggvsyts 552

```

```

RESULT 4
ID AAM37875 standard; Protein; 578 AA.
XX
XX AAM37875;
XX
XX 10-AUG-1998 (first entry)
XX
XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
XX
XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
XX aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
XX 2-keto-L-gulononic acid; L-ascorbic; inhibition.
XX
XX Gluconobacter oxydans.
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note= "signal peptide"
FT 24..578
FT Protein /note= "mature protein"
XX
XX EP832974-A2.
XX
XX 01-APR-1998.
XX
XX 11-SEP-1997; 97EP-0115801.
XX
XX 19-SEP-1996; 96EP-0115001.
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX WPI; 1998-195228/18.
XX N-PSDB; AAV29053.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
XX dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 41-43; 59pp; English.
XX
XX This is the amino acid sequence for the Gluconobacter oxydans
XX alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
XX or recombinant organisms can be used to convert suitable substrates
XX to aldehydes, ketones or carboxylic acids, especially to convert
XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
XX converted to L-ascorbic acid by standard procedures. The derivatives
XX of ADH enzymes have desired substrate specificity, higher affinity
XX to a substrate, lower affinity to an inhibitory compound, higher
XX stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 578 AA:
SQ

```

```

Query Match 85.4%; Score 1736; DB 19; Length 578;
Best Local Similarity 83.8%; Pred. No. 9e-151;
Matches 315; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

```

QY	1	IVANGYIVAGSTCOYSEPFCCFVSGHDSATGGEELMRNYFLPRAGEGDETMGNDYEARMNT	60
Db	179	ivangyivagstcoyspfccfvsghdsatggeelmrnyflpragegdetwngndyearmt	238
QY	61	GAWGQITTDYPTVTLVHYGSTAVGSPASETORGTGGTLLGTWTRPAVRPDTGEIYVRHQTL	120
Db	239	gawgqittdyptvgtlvhygsavgaapaalqrgtvgsswmytntlrtafvpegeiivrinqtl	298
QY	121	PRNMHQECTFEEMATYNDVOSTMEGLQSLNPAAATGERRVLTGVPCKTGTMQPDAAE	180
Db	299	prnmhqectfeemayvndvgsaameglhainpaaatgerrrvltgvpcktgmwqfdae	358
QY	181	TGEFLWARTNTQNNMIESIDENGIYTVNEDAILKLDVEYDVCPTEFGRRDMPGAALNPD	240
Db	359	tgeflwartstqnnliesvdpgjlvhymedlvcelevayelcftlfgyrdwpsaalnpd	418
QY	241	SGITFPLNNVCYDMAADQFTSMQDVYNTSNVTKLPPGKMDIGRIDAIDISTGRTLSVY	300
Db	419	sgitfplnnacsgmtavdqeftssldyynvsldytklpssemgridaidsbgrtlwsa	478
QY	301	ERRAANYSPVLTSGGGVLFNGSTDRYFRALSOETGELTMOGRLATVASSGAISTEYDGMQ	360
Db	479	erryasnyavpvtstggvvlngstdrlyfralsqetgetlwtgrlatvasgaisyeldgvy	538
QY	361	YVAIAGGVSYSGLN 376	
Db	539	yvaigrgrgtsygsnbn 554	
RESULT	5		
AAR05235			
ID	AAR05235	standard; protein; 742 AA.	
XX			
XX	AAR05235;		
DT	04-AUG-1990	(first entry)	
XX			
DE		Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases	
DE		1-2229 of alcohol dehydrogenase (ADH) gene.	
XX			
KM		Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;	
XX		Acetobacter acetii K1006 (FERM-7528); plasmid PAA721.	
OS		Acetobacter acetii strain K1006 (FERM-7528).	
PN		JP02000452-A.	
PD		05-JAN-1990.	
XX			
PF		30-OCT-1987; 87JP-0273190.	
XX			
PR		30-OCT-1987; 87JP-0273190, JP-075069.	
XX			
PA		(QPPP) QP CORP (QPPJ-).	
XX			
DR		WPI; 1990-047990/07.	
DR		N-PSDB; AAQ91811.	
XX			
PT		Cloning DNA, plasmid and microbe contg. it -	
PT		contains alcohol dehydrogenase gene derived from Acetobacter	
XX		acetii K10006 strain	
PS		Disclosure; Fig 4; 8pp; Japanese.	
XX			
CC		Also new are a recombinant plasmid contg. its encoding DNA, and a	
CC		microorganism transformed with the plasmid.	
CC		A DNA fragment was obt'd. from A. acetii	
CC		strain K1006 and was transferred to an ADH-defective strain. The	
CC		resulting plasmid, PAA721, contg. its gene was inserted into A.	
CC		pasteureus strain IFO 3191 by the triparent method. Its gene is	
XX		useful for improving Acetobacter culturing.	

Sequence	742 AA	16.68; Score 338; DB 11; Length 742; Best Local Similarity 26.28; Pred. No. 4,4e-22; Matches 114; Conservative 67; Mismatches 152; Indels 102; Gaps 18;
Query Match		
203 vakgvlvln	---ggaefgargfvsatdaetgkikwrflyvvpnnknepdnaasdnllmka	260
48 -ETGWNDYDARKMT	-----GANGQITRDYDVLNLVHYGSTAVGPASSETORGTRPGTLY	98
261 yktwsp	-kgawvrqg9gglvwdslvydpvsdlly---lavg-----ngspwnkykyse	309
99 -GTN	-----TRAVRPDRDTEIWMRIQTLPDRDMMDOCEFFEMNVTVNDVQPSSTEMGLQSI	152
310 glgslnlfigslvalkpecegywhfgacpmdqdytsvqqnltlompk	-----	358
153 NPNAAITGE	-RRVLTVGPCKTGTMMOFDAETGEFLWARDNTYONMIESID	202
359 -----gemhrviyhap	-kngffyladtgtsfgknvygnwngnjldlqgrmpypnd	411
203 GIATVNEDAILKEIDVEDYVCPTEFGGRWEPALNPDSGITFIPLNNVC	-----	253
412 gilytng	-----kfwygi -pgpialnfmamayspkihlyipahqipfgyknqvggf	463
254 ----DMAVDOEFTSMDYNTSNVTKLPPEKDMIGRIDALIDISTGRTLMSVPRAAANSP	309	
464 kphadswnvgldmtknnglpdipe	-artayikdlhgwllawdpykmetvkiidhkgpntgg	522
310 VLTSGGVLPNGCTRRYRRAALSQETGETLMQTRLATVASGAISIEVDGMOTVA	-----	363
523 ilatgsdillfglgangethaydatngsdlykfdagsgilappmtysvngkqyavavegw	582	
364 ----TACGGVSYSGS	374	
583 gilylsmgvygtrsg	597	
RESULT 6		
AAAR20192		
ID	AAAR20192 standard; Protein; 738 AA.	
AC	AAAR20192;	
DT	16-Apr-1992 (first entry)	
DE	ADH complex protein (mol.wt. 72.000).	
KW	Alcohol dehydrogenase; acetic acid; fermentation.	
OS	Acetobacter altoacetigenes NH-24.	
XX	JP03266988-A.	
XX	27-NOV-1991.	
XX	26-MAR-1990; 90JP-0073440.	
XX	26-FEB-1990; 90JP-0042301.	
XX	26-MAR-1990; 90JP-0073440.	
XX	(NAKA-) NAKANO SUTEN KK.	
XX	WPI; 1992-019325/03.	
XX	N-PSDB; AAQ20383.	
XX	Alcohol dehydrogenase complex structural gene - used in plasmid	
XX	PT and enhancing efficiency of acetic acid fermentation for	
XX	transformed acetic acid bacteria	
XX	Disclosure; Fig 3(1-3); 21pp; Japanese.	

XX Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex
 CC can be easily extracted from the bacteria and purified. It can be
 CC used for the determination of an alcohol.
 CC See also AAQ20383-84, and -86-88.

XX Sequence 738 AA;

Query Match 16.4%; Score 332.5; DB 13; Length 738;
 Best Local Similarity 26.8%; Pred. No. 1.4e-21;
 Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;

QY 2 VANGVIAGSTCYSPFGC--FVSGHDSATGEELMRNYEIPRAGEGD----- 47
 Db 204 lakgrvliqn--gsefsgarqisaftaetgkvdrffvtpnkpnpdaasvlnmkay 261
 QY 48 ETWGNDEYARMTGA-----WCQITVDPYTNLVHGSTAVGPASETGRTGPGTL 97
 Db 262 qtwse-----ptgawtrggsggtvwsivydpradivlylgvngspwnykyrsegkqdn 314
 QY 98 YGTNTRFAVRPDTEIYVRRHQTLPDNDQCEFEAMVTNVDOPTMEGLQSIINPNA 157
 Db 315 lfisgivaikpeltgeywhfgepmdqwdftsqqimldipl----- 357
 QY 158 TGERVVLGVPCRTGTMOQDAETGEFLMARDTNYQNMIESID-ENGIVTVNEDAILKEL 216
 Db 358 ngethrivharngffiyldaklgelisknyvnmwasgldpkrpilynpdalytl 417
 QY 217 DVEYDVCPTFLGGRDMSALNPDGSIYFIPLNVCYDMAVDOEFT-SMDVYNTS---N 272
 Db 418 gkewygiipgdilgghnfamaafspkgtlylpaqgvplytnqvgfipbdswnlgldm 477
 QY 273 VTKLPPG-----KDMIGRIDAIDISTGRTLMSVERAANSPVLSTGGVLFNGTD 324
 Db 478 kvgiipdspeakqatfvkdkyivawdpqkgeaavrvdhkpgmngilatgdlifgqlan 537
 QY 325 RYFRALSOETGETLMQTRLATVASGQAISEYVDMQYVAIAGG 367
 Db 538 gefhaydatngsdlfhfaadsgliappvtylangkgyvavavg 580

RESULT 7

AA13993
 ID AAR13993 standard; Protein; 738 AA.

AA13993;

09-DEC-1991 (first entry)

A.altoacetigenes membrane-bound ADH 72KD sub-unit.

alcohol dehydrogenase complex; carboxylic acid production.

Acetobacter altoacetigenes.

EP448969-A.

02-OCT-1991.

26-FEB-1991; 91EP-0102793.

26-MAR-1990; 90JP-0073440.

26-FEB-1990; 90JP-0042391.

(NAKA-) NAKANO VINEGAR KK.

Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;

WPI: 1991-289462/40.

N-PSDB; AAQ13580.

XX Gene for membrane-bound alcohol dehydrogenase complex - obtd.
 PT from Acetobacter altoacetigenes, used for prodn. of enzyme for
 PT converting alcohol to acid

PS Disclosure: Fig 3; 36pp; English.

XX Total DNA was prepared from A.altoacetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
 CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A.altoacetigenes (see AAQ13582-013584). The gene
 CC encoding the 72KD ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72KD sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72KD
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.

XX Sequence 738 AA;

Query Match 16.3%; Score 331.5; DB 12; Length 738;
 Best Local Similarity 26.8%; Pred. No. 1.7e-21;
 Matches 108; Conservative 52; Mismatches 180; Indels 63; Gaps 10;

QY 2 VANGVIAGSTCYSPFGC--FVSGHDSATGEELMRNYEIPRAGEGD----- 47
 Db 204 lakgrvliqn--gsefsgarqisaftaetgkvdrffvtpnkpnpdaasvlnmkay 261
 QY 48 ETWGNDEYARMTGA-----WCQITVDPYTNLVHGSTAVGPASETGRTGPGTL 97
 Db 262 qtwse-----ptgawtrggsggtvwsivydpradivlylgvngspwnykyrsegkqdn 314
 QY 98 YGTNTRFAVRPDTEIYVRRHQTLPDNDQCEFEAMVTNVDOPTMEGLQSIINPNA 157
 Db 315 lfisgivaikpeltgeywhfgepmdqwdftsqqimldipl----- 357
 QY 158 TGERVVLGVPCRTGTMOQDAETGEFLMARDTNYQNMIESID-ENGIVTVNEDAILKEL 216
 Db 358 ngethrivharngffiyldantgetisgknyvnmwasgldpkrpilynpdalytl 417
 QY 217 DVEYDVCPTFLGGRDMSALNPDGSIYFIPLNVCYDMAVDOEFT-SMDVYNTS---N 272
 Db 418 gkewygiipgdilgghnfamaafspkgtlylpaqgvplytnqvgfipbdswnlgldm 477
 QY 273 VTKLPPG-----KDMIGRIDAIDISTGRTLMSVERAANSPVLSTGGVLFNGTD 324
 Db 478 kvgiipdspeakqatfvkdkyivawdpqkgeaavrvdhkpgmngilatgdlifgqlan 537
 QY 325 RYFRALSOETGETLMQTRLATVASGQAISEYVDMQYVAIAGG 367
 Db 538 gefhaydatngsdlfhfaadsgliappvtylangkgyvavavg 580

RESULT 8

AAB35987
 ID AAB35987 standard; Protein; 754 AA.

AAB35987;

01-MAR-2001 (first entry)

Sorbitol dehydrogenase subunit 1 amino acid sequence.

Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;

L-sorbitol production; 2-keto-L-gulonic acid.

Gluconobacter oxydans.

WO200065066-A1.

PD 02-NOV-2000.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX (CHOI/) CHOI E.
 XX (RHEE/) RHEE S.
 XX (LEE/) LEE E.
 XX
 XX Choi E, Rhee S, Lee E;
 XX
 XX MPI: 2000-687351/67.
 XX N-PSDB: AAC83153.
 XX
 XX Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 XX isolated from Gluconobacter suboxydans useful for the fermentative
 XX production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol
 XX
 XX Claim 1: Fig 8; 96pp; English.
 XX
 XX This invention relates to an isolated membrane-bound sorbitol
 XX dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 XX polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 XX (AA835987 - AA835989). Also included in the invention are two
 XX polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 XX coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 XX useful for producing L-sorbose from D-sorbitol and for increasing the
 XX production of 2-keto-L-gulonic acid by transforming a host cell,
 XX especially Gluconobacter with the DNA and selecting the transformed host
 XX cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 XX genes) encoding fragments of SDH are specifically claimed, however these
 XX sequences are not given separately in the specification but are included
 XX in sequences AAC83156 and AAC83157.
 XX
 XX Sequence 754 AA:

Query Match 16.1%; Score 326.5; DB 21; Length 754;
 Best Local Similarity 25.3%; Pred. No. 5.1e-21;
 Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGYIVAGSTQCYSPFC--FVSGHDSATGEBELMRNYFIRAGEEGD-----47
 DB 199 jakgvlign--ggaeifargftaydaetckmdwrtfvtvnpdkpdaasddvlnska 256
 QY 48 -ETWGNDEARWMTG--AMGQITVDPVTLNHYGSTAVGPASETORGTGPGTLYGTWTR 103
 DB 257 yltwkggawkgqggggtvwdsllydvtldlylgvgngspwnykftiseqgnmlfigsl 316
 QY 104 FAVRPDGEIIVRHOTLPDRDWMDOCTEFEMAVTNVDQPSSTEMEGLOSINPNAATGE-RR 162
 DB 317 valnpdtkywhfgetpmdgdwlytsvqqlmalmpv-----ngemth 359
 QY 163 VLTGVPCKTGTMQFDAGTGEFLWADNTNYNMIESIDE-NGIYTVNEDATLKELDVEYD 221
 DB 360 vlnhap-kngftlyldatgkfytsyngwngldpvtgrpnypdaiwlclngkpwpy 418
 QY 222 VCPFFLGSRMPSAALNDGSIYPIPLANNCY-----DMKAVDOETSMVY 268
 DB 419 gipddlgghmtaamayspqtilyipaqvpyfvdpqkggfkahhdswnlglldmnc 478
 QY 269 NTSN---VTLPPGKDMIGRIDIDISTGRTLSVERAANSPVLTSGGVLFNGGTD 324
 DB 479 ddnopkhkadagfikldkgywvawdpqkgaatlvdkhkpwnqllatagvylfgqlan 538
 QY 325 RYFRALSOETGETLMQTRLATVASGOAISYEVDQMOYVAI---AG-----GVSVSGS 374
 DB 539 gefhayattgkdlftfpagaaiaipvtylangkgyvavewgvgilypiffigvarts 598

RESULT 9
 AAW95019

ID AAW95019 standard; Protein; 740 AA.
 XX
 XX AAW95019;
 XX
 XX 21-MAY-1999 (first entry)
 XX
 XX Sorbitol dehydrogenase (SDH) protein sequence.
 XX
 XX Sorbitol dehydrogenase; SDH; open reading frame; ORF2; L-sorbose;
 XX D-sorbitol; vitamin C; enzyme.
 XX
 XX Gluconobacter suboxydans.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..24
 XX Protein /note="signal sequence"
 XX /note="740
 XX /note="mature protein"
 XX
 XX EP897984-A2.
 XX
 XX 24-FEB-1999.
 XX
 XX 13-AUG-1998; 98EP-0115231.
 XX
 XX 21-AUG-1997; 97EP-0114432.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 XX Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;
 XX MPI: 1999-134646/12.
 XX N-PSDB: AAX21501.
 XX
 XX New D-sorbitol dehydrogenase gene and recombinant protein - useful
 XX for production of L-sorbose, an intermediate in vitamin C production
 XX
 XX Claim 1: Fig 3a-D; 39pp; English.
 XX
 XX This represents a sorbitol dehydrogenase (SDH) protein. The DNA
 XX encoding the SDH enzyme also encodes an open reading frame (ORF2)
 XX product upstream of the SDH open reading frame, needed for SDH activity
 XX in vivo. Host cells transformed by a vector comprising the sorbitol
 XX dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 XX producing L-sorbose from D-sorbitol. L-sorbose is an important
 XX intermediate in vitamin C production.
 XX
 XX Sequence 740 AA:

Query Match 9.9%; Score 202; DB 20; Length 740;
 Best Local Similarity 21.2%; Pred. No. 1.4e-09;
 Matches 108; Conservative 51; Mismatches 144; Indels 206; Gaps 24;

QY 2 VANGYIVAGSTC-----QYSPFCFVSGHDSATGEBELMRNYFIRAGEEGDFTWGNDEYA 56
 DB 273 vlnqvvvnnhevlldggrtwapsq-vlrgydaesgfw-----awdvmsn 317
 QY 57 RMTGTAMGQITVDPVTLNHYGSTAVGPASETORGTGPGTLYGTWTR-----RAY 106
 DB 318 r-----sqpay-rvvlaverrlpjldrtirg-----gsrlrpdnsaadysal 363
 QY 107 RPD-----TGEIIVRHOTLPDRDWMDOCTEFEMAVTNVDQPSSTEMEGLOS 151
 DB 364 rsdaenksavvaldvktgspvrfqctahkdvwdydgatli-----mdmpg--- 412
 QY 152 INPNAATGERRVLTGVPCKTGTMQFDAGTGEFLWADNTNYNMIESIDE-----GIV- 205
 DB 413 -pdgqtlvpallm--ptkrqgtfvldrrtgk-----pdlpveeapapapyldp 456
 QY 206 -----TYNEDAI-----LKEID----- 217

Db 457 gdprrptgswvmpalrypdlkeldmwmgspldqlfcrifiranygvefpevdkpw 516
 QY 218 VEDVOCPTFLGGDRPMSALNDPDSGIFRPLN-NVCYDMAVDOFTSM----- 265
 Db 517 ley---pyngsgdwsmysdpqsglllanwiltmrydltvrkxadsiglmiddpnfk 573
 QY 266 -----DVTNTSNVTKLPKCKMIGRIDIDISTG-RTIMS 299
 Db 574 pgggaagngandgtpyglvtrpfndqy-tgmmcnrpp---ymtaldmkgqkvlwq 628
 QY 300 -----VERAANYSPVLSTGGGVLFNG-GTDRYPRALSOETGTLW 339
 Db 629 hplgtarngpwlplglpwltpngsgvvtgg9llffigaatnglraidehtgkvww 688
 QY 340 QTRLATVASGQASIVYVDCMNOVATAGG 368
 Db 689 savlpgggaanpmtyleanghqvaylmaag 717

RESULT 10

AAU12236 standard; Protein; 530 AA.

AAU12236;

24-OCT-2001 (first entry)

Human PRO4340 polypeptide sequence.

Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 adipocyte; A-peptide; factor VIIA; gene therapy.

Homo sapiens.

WO200140466-A2.

07-JUN-2001.

01-DEC-2000; 2000MO-US32678.

01-DEC-1999; 99MO-US28301.
 01-DEC-1999; 99MO-US28634.
 02-DEC-1999; 99MO-US28551.
 02-DEC-1999; 99MO-US28564.
 02-DEC-1999; 99MO-US28565.
 09-DEC-1999; 99MO-US30095.
 16-DEC-1999; 99MO-US30911.
 20-DEC-1999; 99MO-US30999.
 30-DEC-1999; 99MO-US31243.
 06-JAN-2000; 2000MO-US00277.
 06-JAN-2000; 2000MO-US00376.
 11-FEB-2000; 2000MO-US03365.
 18-FEB-2000; 2000MO-US04341.
 22-FEB-2000; 2000MO-US04414.
 24-FEB-2000; 2000MO-US04914.
 24-FEB-2000; 2000MO-US05004.
 01-MAR-2000; 2000MO-US05601.
 20-MAR-2000; 2000MO-US07377.
 21-MAR-2000; 2000MO-US07532.
 30-MAR-2000; 2000MO-US08439.
 17-MAY-2000; 2000MO-US13705.
 22-MAY-2000; 2000MO-US14042.
 30-MAY-2000; 2000MO-US14941.
 02-JUN-2000; 2000MO-US15264.
 10-NOV-2000; 2000MO-US30873.

(GETH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;

PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI: 2001-408281/43.
 DR N-PSDB: AAS21308.

PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical

Claim 12; Fig 130; 813pp; English.

AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, the polynucleotides encoding
 CC transgenic or knock out animals and can be used in gene therapy.

Sequence 530 AA;

Query Match 5.3%; Score 108.5; DB 22; Length 530;

Best Local Similarity 24.3%; Pred. No. 0.33; Matches 79; Conservative 38; Mismatches 109; Indels 99; Gaps 20;

QY 59 MTGANGQITTP-VTNLVHVGSTAVGPAERORIGPGTL-----YGTNTR--- 103
 Db 18 isgcwgvnrllpfiftn--hfdtvlisedtprvsgsvtqlagdmddpilyfsgeas 75
 QY 104 --FAVRPTGETIWMRHOTLPDNDQECTFEAMVYN-----VDVOPSTEMGLASIN 153
 Db 76 rffavepdyg-vwvllrpldret-kseftvefsdngqvltirvnlq-----vgdvn 126
 QY 154 PNAATGER-----RYLVGVPKT-----GTMOPDAETGEFLNARDTNY 192
 Db 127 dnaptfhngpysvripentpvgtpflivnatcdpdlgaagsvlysfqpsqf----- 178
 QY 193 QNMISIDE-NCIYVNVNDALIKELDVE-----YDPCPFLGGRMPMSALNDPDSGIFR 247
 Db 179 ----aisargivlv-----ireldyeltqayqltvn-----atqdktrplst 218
 QY 248 INN--VCYDMAVDOFTSMVNTSNVTKLPKCKMIGRIDIDISTGRTL--MSVE 301
 Db 219 lanlaillitdvgdmddpiflnp-ystlnyehsppp-ttvlrlltaldqkgypgijytlv 276
 QY 302 RAAANYSPVLSTGGGVLFNGGTR 325
 Db 277 sgntsfaldyisgvltnlglldr 301

RESULT 11

AAU07006 standard; protein; 269 AA;

AAU07006;

17-JAN-1991 (first entry)

[illegible][illegible]

Db 212 rg--daptepekeasplvp1 231

RESULT 13

AA07014 standard; protein; 593 AA.

AA07014;

17-JAN-1991 (first entry)

Protein G variant with three active sites.

Immunoglobulin.

Streptococcus sp. Lancefield Group G strain.

Key Location/Qualifiers

Active-site 303..372 /label=B1

Active-site 373..427 /label=B3

Active-site 443..497 /label=B2

US4956296-A.

11-SEP-1990.

20-JUN-1988; 88US-0209236.

20-JUN-1988; 88US-0209236.

14-FEB-1986; 86US-0829354.

23-APR-1986; 86US-0854887.

17-FEB-1987; 87MO-US00329.

19-JUN-1987; 87US-0063959.

(GENE-) GENEX CORP.

Fahnestock SR.

WPI: 1990-297491/39.

N-PSDB; AA006019.

Recombinant Protein G variants - obid. using a cloned gene encoding Protein G from Streptococcus sp., used for binding immunoglobulin.

Disclosure; Fig 9; 48pp; English.

Fragments and variants of the sequence are claimed esp. where incorporated into a non-pathogenic host eg. E.coli, and expressed at high levels.

The variants have a higher binding efficiency and capacity for immunoglobulin, and may be used for purifying, detecting and isolating antibodies.

Sequence 593 AA;

Query Match 5.2%; Score 106.5; DB 11; Length 593;

Best Local Similarity 25.4%; Pred. No. 0.6; Indels 55; Gaps 14;

Matches 66; Conservative 32; Mismatches 107; Indels 55; Gaps 14;

1 IVANGVIVAGSTCOYSPFGVSGHDSATGELMKNYFIRPAGEGSDTWCNDYEARMWT 60

306 llnngtlkgetl-----teavdaataekvfky-----and-----n 338

61 GAWGQITTPDPVTLVH-----YGSTAVGPASETORGTGG-TLYGINTFRVAPPDIGE 112

339 gvdgvcwtydaktftvtekevidaseltptvtykpyingklklygetlteavdaatae 398

113 IYVRHQTLPFRDN-WDEGCTFEMVNTVNDQPSSTEMEGLOSINPMNATGERRVLTGVPCKT 171

Db 399 kvfkqya--ndhvvdgewctydatkftvtekevidaselp-avtyklyvngklkg 455

172 GTMMQ-FDAETGELWARPNTYONMIESIDENGI--VIVNEDAILKELDVEYDVPTEFG 228

456 elttkavdaetaekafkqyan-----dnvgdvwvtydda-ekftvtemv---tevp 503

229 GRDWPSSALNPDGIVETPL 248

504 g-daptepekeasplvp1 522

RESULT 14

AA010005 standard; protein; 594 AA.

AA010005;

13-MAR-1991 (first entry)

Streptococcus GX7805 protein G.

Immunoglobulins; Ig.

Streptococcus sp GX7805.

Key Location/Qualifiers

Active-site 304..358 /label= Active Site B1

Active-site 374..428 /label= Active Site B3

Active-site 444..498 /label= Active Site B2

US4977247-A.

11-DEC-1990.

19-MAY-1989; 89US-0354264.

19-MAY-1989; 89US-0354264.

14-FEB-1986; 86US-0829354.

23-APR-1986; 86US-0854887.

17-FEB-1987; 87MO-US00329.

19-JUN-1987; 87US-0063959.

20-JUN-1988; 88US-0209236.

(GENE-) GENEX CORP.

Fahnestock SR, Lee T, Wroble MH;

WPI: 1991-006758/01.

O-PSDB; Q10002.

Immobilised protein G variants - used for detection, isolation and purification. Immunoglobulin(s) and immunoglobulin fragments

Disclosure; Fig 9; 52pp; English.

Protein G gene product may be modified allowing the variant to be immobilised and exhibit different binding profiles. The bound protein is useful in purification and detection of Igs and fragments.

Sequence 594 AA;

Query Match 5.2%; Score 106.5; DB 12; Length 594;

Best Local Similarity 25.4%; Pred. No. 0.6; Indels 55; Gaps 14;

Matches 66; Conservative 32; Mismatches 107; Indels 55; Gaps 14;

1 IVANGVIVAGSTCOYSPFGVSGHDSATGELMKNYFIRPAGEGSDTWCNDYEARMWT 60

307 llnngtlkgetl-----teavdaataekvfky-----and-----n 339

```

QY 61 GAMGQITVDYPTNLVH-----YGTAVGPASERGTGPGG-TLYGTMTREAVRPDGE 112
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 340 gvdgwyddatktftvtekpevidaseltpavtlkpylingkllggetttaevdaetae 399
QY 113 IVMRHQTLPRDN-WDOECTFEMAVTNVDVQPTSTEMEGLQSIINPNATGERRVLTGVPCKT 171
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 400 kvtkqya--ndnvgdewlyddatktftvtekpevidaseltp-avtlkklivingkllkg 456
QY 172 GTMMQ-FDAETGELFMDRTNNTQNMIESIDENGI--VTNEDAILKELDEYDVCPTFG 228
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 457 ettkaevdaetaekafkqyan-----dngvdgvywlydda-ktftvtemv--tevp 504
QY 229 GRDMPASALNPDGTYFIPL 248
Db 505 g-dapekpeasipvl 523

```

RESULT 15

AAP94788
ID AAP94788 standard; protein; 269 AA.

AC AAP94788;

DT 04-JUL-1990 (first entry)

DE Protein G variant.

KW Protein G; immunoglobulin; Fc receptor; ds.

OS Streptococcus sp.

PN W08810306-A.

PD 29-DEC-1988.

PF 20-JUN-1988; 88MO-US02084.

PR 19-JUN-1987; 87US-0063959.

PA (GENE-) GENEX CORP.

PI Fahnstock SR;

DR WPI; 1989-023848/03.

DR N-PSDB; AAN94674.

PT Cloned protein G variant genes -
expressing proteins having immunoglobulin-binding properties of
protein G and derived from Streptococcus sp.

PS Clalm 19; Page 89; 116pp; English.

CC Gene for protein G variant of non-pathogenic streptococcus sp. allowing
isolation of the protein and variants, useful as bacterial Fc receptors
CC eg in purification and detection of Abs., screening of hybridoma clones
CC and treatment of disease.

XX Sequence 269 AA;

Query Match

Best Local Similarity 5.2%; Score 106; DB 10; Length 269;
Matches 64; Conservative 30; Mismatches 110; Indels 58; Gaps 13;

```

QY 1 IVANVIVAGSTCGSPGCFVSGHDSATGELMRNRYIPRAGEGDETMNDYEARMT 60
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 lllngkllggett-----teavdaetaekviky-----and-----n 46
QY 61 GAMGQITVDYPTNLVHVSSTAVGPASERGTGPGG-TLYGTMTREAVRPDPT 110
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 gvdgwyddatkt--ftvtekpevidaseltpavtlkklivingkllggetttaevdaet 104

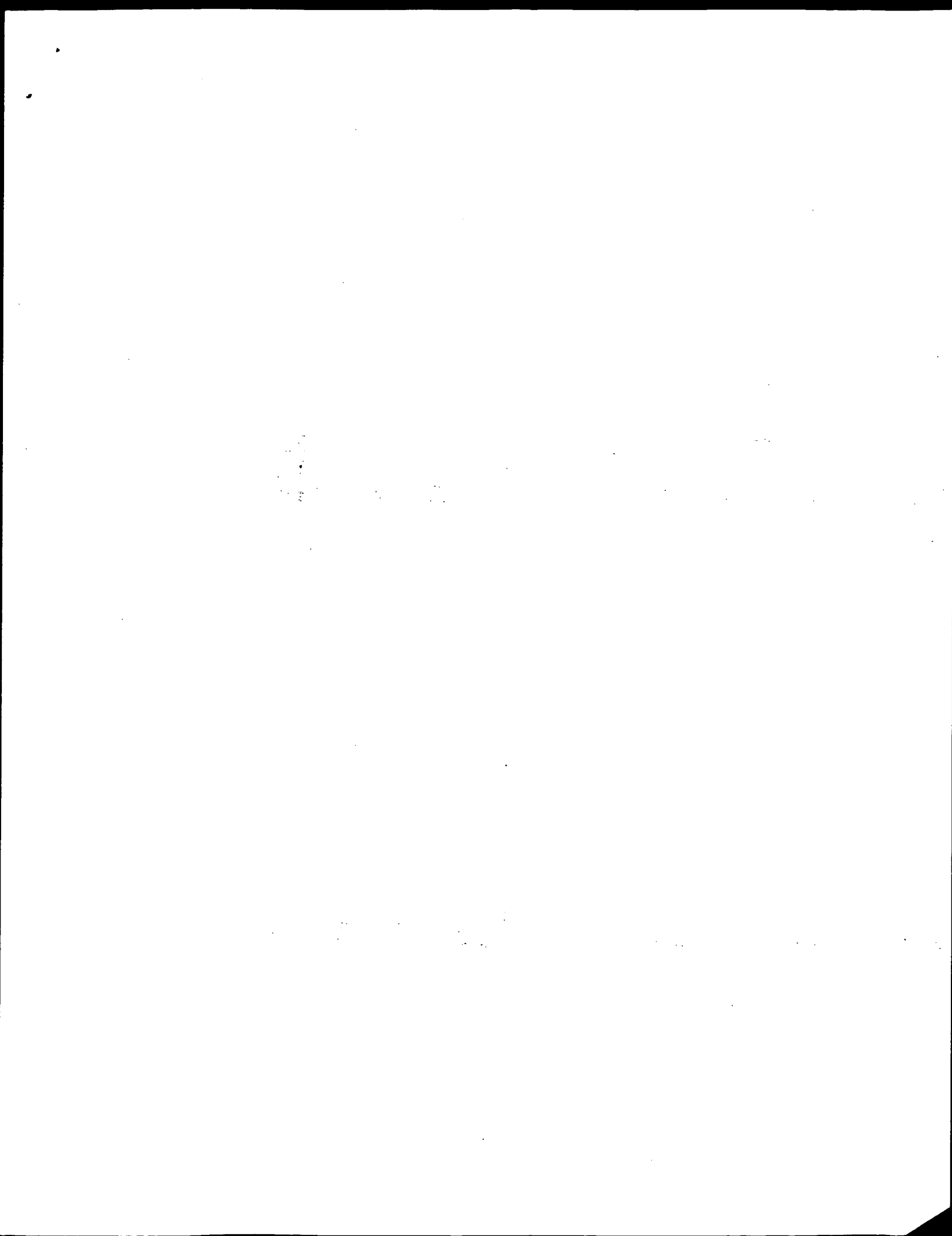
```

```

QY 111 GEIVMRHQLPRDNW-DOECTFEMAVTNVDVQPTSTEMEGLQSIINPNATGERRVLTGVPCK 169
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 105 aekvtkqya--ndnvgdewlyddatktftvtekpevidaseltp-avtlkklivingkll 161
QY 170 GTMMQ-FDAETGELFMDRTNNTQNMIESIDENGI--VTNEDAILKELDEYDVCPTFG 226
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 162 kgettaevdaetaekafkqyan-----dngvdgvywlydda-ktftvtemv--tevp 211
QY 227 LGGRDMPASALNPDGTYFIPL 248
Db 212 rg--dapekpeasipvl 231

```

Search completed: May 24, 2002, 10:05:04
Job time: 243 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:02:58 ; Search time 44.35 Seconds

(without alignments)
207.631 Million cell updates/sec

Title: US-08-934-506A-5_COPY_180_556
Perfected score: 2032
Sequence: 1 IVANGYIVAGSTGQYSPFGC.....GMQYVAIAGGQVSYGSLNS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.5	16.4	738	1	US-07-985-458-3
2	326.5	16.1	720	4	US-09-296-284-25
3	326.5	16.1	754	4	US-09-296-284-4
4	207	10.2	739	3	US-09-136-251-2
5	105	5.2	1012	1	US-08-219-262B-10
6	105	5.2	1012	3	US-09-031-655-10
7	95.5	4.7	548	5	PCT-US93-10541-2
8	95	4.7	1012	1	US-08-216-276A-19
9	94.5	4.7	512	4	US-09-356-818A-2
10	93.5	4.6	548	1	US-08-247-902A-2
11	93	4.6	380	4	US-08-971-782-4
12	93	4.6	380	4	US-09-309-026-4
13	93	4.6	459	4	US-08-971-782-2
14	93	4.6	459	4	US-09-309-026-2
15	91.5	4.5	606	2	US-08-883-534-3
16	91.5	4.5	606	3	US-09-204-764-3
17	91.5	4.5	816	1	US-07-731-157A-4
18	91.5	4.5	816	1	US-08-229-444B-2
19	91.5	4.5	816	2	US-08-541-780-4
20	91.5	4.5	1612	1	US-08-169-927-2
21	91	4.5	551	2	US-08-193-229-32
22	91	4.5	551	3	US-09-285-957-32
23	90.5	4.5	322	1	US-08-216-276A-29
24	90.5	4.5	687	5	PCT-US91-09784-2
25	90	4.4	484	1	US-08-216-276A-17
26	90	4.4	1012	1	US-08-219-262B-1
27	90	4.4	1012	1	US-08-219-262B-12

28	90	4.4	1012	3	US-09-031-655-1	Sequence 1, Appl
29	90	4.4	1012	2	US-09-031-655-12	Sequence 12, Appl
30	90	4.4	1222	3	US-08-682-517-15	Sequence 9, Appl
31	90	4.4	1252	2	US-08-682-517-9	Sequence 2, Appl
32	88	4.3	1012	1	US-07-944-943-2	Sequence 2, Appl
33	88	4.3	1012	1	US-08-219-262B-2	Sequence 2, Appl
34	88	4.3	1012	3	US-09-031-655-2	Sequence 2, Appl
35	88	4.3	1287	1	US-08-200-232-2	Sequence 2, Appl
36	88	4.3	1287	5	PCT-US95-02219A-2	Sequence 2, Appl
37	88	4.3	1287	5	PCT-US95-02219A-2	Sequence 2, Appl
38	87.5	4.3	592	1	US-08-217-327-8	Sequence 8, Appl
39	87.5	4.3	635	4	US-08-931-608A-5	Sequence 8, Appl
40	87	4.3	395	4	US-08-856-841-13	Sequence 13, Appl
41	87	4.3	395	4	US-08-856-841-22	Sequence 22, Appl
42	85.5	4.2	723	1	US-08-290-937B-3	Sequence 3, Appl
43	85	4.2	774	4	US-09-346-237-8	Sequence 8, Appl
44	85	4.2	5215	4	US-09-105-537-2	Sequence 2, Appl
45	84	4.1	500	2	US-08-987-519-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-985-458-3
Sequence 3, Application US/07985458
Patent No. 5344777
GENERAL INFORMATION:
APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takemura, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hejima and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
NUMBER OF INVENTION: 4
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fishauf, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about
IDENTIFICATION METHOD: 72,000
ORGANISM: Acetobacter alioacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Tayama, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horinouchi, Sueharu and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

Query Match 16.4%; Score 332.5; DB 1; Length 738;
Best Local Similarity 26.8%; Pred. No. 5,7e-25;
Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47
DB 204 IAKGVIIIGN--GGSEFGARGFVSAPDAETGKVDNRFETPRPKNEPDASDVLNKKAY 261
QY 48 ETWGNDEYEAHMMTG-----WGQITTPVTNLVHYGSTAVGPASSTORGTGTNTR 97
DB 262 QTWS-----PTGAWTRGGGGGTWDSIVYDPVADLVYLGVNGSPMNYKRSSEKGDN 314
QY 98 YGTFEFAVRDPGEIYWRHOTLPKRDNDQECTFEMMTYNVVOPTSTEMGLQSLNPMNA 157
DB 315 LPLGSLVALKPTGETGVVHFQETPRDQWDFTSDDQIMTLDFI----- 357
QY 158 TGERVRLGVPCKTGTMMQFDETEGEFLMARDTNTQNMIESID-ENGIVTVEDAILKEL 216
DB 358 NGETRHYIYHARKNGFPYIDAKTGEFISGKNYVYVNNASGLDPRGTINPDLTYLT 417
QY 217 DVEYDVCPTFLGDRMPKSAALNPDSCGIFIPLNVCYMMMAVDQET--SMDVYMTS--N 272
DB 418 GKEMWGIPDLGHNFAAMASPKTGLVYIPAQVPEFLYTVQVGGFTPHDPSMNLGIDMN 477
QY 273 VKLPPG-----KDMIGRIDIDISTGRTLMSVERAANYSPVLSTGGGVLFNGSTD 324
DB 478 KVGIDPSEAKQAFYKDKLGMIVAMPDPOKQAEAMRVDDKGPWNGGILATGDLFQGLAN 537
QY 325 KYFRLALSGEETLMQRLATVVASGQALSYEVDSQYVAIAG 367
DB 538 GEFHAYDATNGSDLFHFAADSGIATPPTIYLANCKQYVAIVEG 580

RESULT 2
US-09-296-284-25
Sequence 25, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Gluconobacter suboxydans sorbitol dehydrogenase, Genes

TITLE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
US-09-296-284-25

Query Match 16.1%; Score 326.5; DB 4; Length 720;
Best Local Similarity 25.5%; Pred. No. 2,4e-24;
Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47
DB 165 IAKGVIIIGN--GGSEFGARGFVSAPDAETGKVDNRFETPRPKNEPDASDVLNKKAY 261
QY 48 ETWGNDEYEAHMMTG-----WGQITTPVTNLVHYGSTAVGPASSTORGTGTNTR 97
DB 223 YPTMGKGAAMKQGGGGGTWDSIVYDPVADLVYLGVNGSPMNYKRSSEKGDN 314
QY 104 FAVRPDGEIYWRHOTLPKRDNDQECTFEMMTYNVVOPTSTEMGLQSLNPMNA 157
DB 283 VAINDDTGKVVWFQETPRDQWDFTSDDQIMTLDFI----- 357
QY 163 VLTGVPCKTGTMMQFDETEGEFLMARDTNTQNMIESID-NGIYTVEDAILKELDYED 221
DB 326 VLVNAP--KNGEYIIDLKTKFKISGKRYEENWANGLDVPTGRPNYDPLMTLNGKPMY 384
QY 222 VCPFLGDRMPKSAALNPDSCGIFIPLNVCY-----DMMAVDQETSMY 268
DB 385 GIPDLGSHNFAAMASPKTGLVYIPAQVPEFLYTVQVGGFTPHDPSMNLGIDMN 477
QY 269 NTSN-----VTKLPPGKMGIRIDIDISTGRTLMSVERAANYSPVLSTGGGVLFNGSTD 324
DB 445 DDNDPOHKAADKAPFLKDKLGMIVAMPDPOKQAEAMRVDDKGPWNGGILATGDLFQGLAN 537
QY 325 KYFRLALSGEETLMQRLATVVASGQALSYEVDSQYVAIAG 367
DB 505 GEFHAYDATNGSDLFHFAADSGIATPPTIYLANCKQYVAIVEG 580

RESULT 3
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Gluconobacter suboxydans sorbitol dehydrogenase, Genes
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
US-09-296-284-4

Query Match 16.1%; Score 326.5; DB 4; Length 754;
Best Local Similarity 25.5%; Pred. No. 2,4e-24;
Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47


```

Db 199 IAKGVIITN--GGAEFGARGVITAYDFTGKMDKREFTTVNPNPKPPGASDDVLSKA 256
QY 48 -ETMGNDYEAARMUTG---AMGQITYPVTLNVHGSTAVGPASETORCTPGCTIGTNTFR 103
Db 257 YPTWCKGAMKQGGGGGYMDSLIYDPTDLYLGVGNGSPWNYEFRESEGNMFLGSI 316
QY 104 FAVRPDTGEIYWRHQTLPDRDNMDQECTEMAVTNDVQPSIEMELQSIINNAATGE-RR 162
Db 317 VAINPDGTGYVWHFQETPMDQMDYTSVOQINALMDPV-----NGENRH 359
QY 163 VLTGVPCKTGMQDFAETGEFLMARDTNYONMIESIDE-NGIYVNDAILKEIDVEYD 221
Db 360 VLVHAP-KNGEFTYIIDAKTGKTSKPTTYENWANGLDPRGPRVYNPDALMTLNGKPMY 418
QY 222 VCPTELGHRDMSALNPDGSIYFIPLNNVCY-----DMAVDOFTSMQY 268
Db 419 GJPGDLGHNFAAMVSPQTKLYIPAOQVPFVYDPOKGFKAHDSWNLGIDMKKICIL 478
QY 269 NTSN-----VTKLPKCKDMIGRIDALDISTGRILMSVERAANYSPVLSTGGVLFNGGTD 324
Db 479 DUNDPOHKADKQFLKDLKGMIVAMPDQKQAFVVDHKGPNNGGLATAGVLFQGLAN 538
QY 325 RYFRALSOETGELMOTRLATVAGSAISYEVDGMQYVAI---AG-----GVSYSGS 374
Db 539 GEFHATVDTGKDLFTFPQASALIAIPVITYTANGKQYVAVEVGMGIVPFLGVAARTSG 598

```

```

RESULT 4
US-09-136-251-2
; Sequence 2, Application US/09136251A
; Patent No. 6127156
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: MIYAZAKI, Taro
; APPLICANT: OJIMA, Setsuko
; APPLICANT: SHINOH, Masako
; APPLICANT: TOMIYAMA, No. 61271561bun1
; TITLE OF INVENTION: D-Sorbitol dehydrogenase gene
; FILE REFERENCE: D-Sorbitol dehydrogenase gene
; CURRENT APPLICATION NUMBER: US/09/136,251A
; EARLIER APPLICATION NUMBER: EP 97114432.4
; EARLIER FILING DATE: 1997-08-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(24)
US-09-136-251-2

```

```

Query Match 10.2%; Score 207; DB 3; Length 739;
Best local similarity 21.5%; Pred. No. 3, le-12;
Matches 109; Conservative 51; Mismatches 142; Indels 206; Gaps 24;

```

```

QY 2 VANGVIYAGSTC-----QISPFQCFVSGHSAANGELIMRYFIIPRAGEBDETWGNDYEA 56
Db 273 VINGVVVNEHVLIDGQRRMAPSG-VIRGYDAESGKFEV-----AMDVNSG 317
QY 57 RMMTGAGQITYPVTLNVHGSTAVGPASETORCTPGCTIGTNTNR-----FAV 106
Db 318 R-----SQPAY-RVYITAVEKRIPLPRRQSG-----GSRRLRPDMSAADYLSL 363
QY 107 RPD-----TGEIWRHQTLPDRDNMDQECTEMAVTNDVQPSIEMELQSI 151
Db 364 RSDAENKVSAAVAIDVKTGSPRVFQTAHRDWDYDIGSQATL-----MDMG--- 412
QY 152 INPNAATGERRVLTGVPCKTGMQDFAETGEFLMARDTNYONMIESIDEN-----GIY- 205

```

```

Db 413 --PDGQVFPALIM---FTKRGQTFVLDRTGK-----PILPVEERAPSPGVIIP 456
QY 206 -----TVNEDAI-----LIKED----- 217
Db 457 GDBRSPPTQPSVGMFALRPVDELKETDWMGMSPIDQLECFKIFRANVGEFTPPSVDKPW 516
QY 218 VEYDVCPTELGHRDMSALNPDGSIYFIPLN-NVCYDMAVDOFTSM----- 265
Db 517 IEY---PGYNGSGPWSGMSYDPOSGILIANMNITPMYDQLVTRKKADSLGIMPIDPNFK 573
QY 266 -----DYNTSNVTKLPKCKDMIGRIDALDISTG-RTIWS 299
Db 574 PGCGAGAGCAGMDGTPYGIWVTFPMDQY-TGMNCRPP-----YGMTAIDMKHGQVLMQ 628
QY 300 -----VERAANYSPVLSTGGGVLFNG-GTDYFRALSOETGELM 339
Db 629 HPLGTARANGFWGLPGLPWFELGTPNNGSVYVGGGLIFIGATDNGIRAIIDHTGKVVW 688
QY 340 QTRLATVAGSAISYEVDGMQYVAIAGG 367
Db 689 SAVLPGGQANPMTYEANGHGYVAIAGG 716

```

```

RESULT 5
US-08-219-262B-10
; Sequence 10, Application US/08219262B
; Patent No. 5786970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNIDER, DAVID B
; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: OBLON, SPIYAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 175 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: OH
US-08-219-262B-10

```

Query Match 5.2%; Score 105; DB 1; Length 1012;
Best Local Similarity 22.2%; Pred. No. 0.12;
Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 71 VTNLVHGSTAV-----GPAS-----ETQGTGGTLYGTNTRF 104
DB 1 MTNLMHTQOIVPFRSLMPTGPAISIPDDTLEKHLRSETSYNLTVGDTGSGLIVF 60
QY 105 AVRPDGEIVMRHQTLPD---NMDOECTFEMMT-----NDVQFSTE 145
DB 61 PGFP--GSVGAHHTLOSNGSYQFDQ-----MLTQNLPSVSYNCRVSRSLTVRSSTL 113
QY 146 MEGLOSINP--NAATGERRVLTGVPCKTGMQFDETFELMARDTNYQNMIESI---- 199
DB 114 PGVYALNGTINAVT-----FGSLSEL---TDYSNGMSATANIN 152
QY 200 DENGIVYVNEALIKELDEYDVCPTFLGGRDPSAALNDPSGIYFIPLNNVC----- 252
DB 153 DKIGNVTVGEGVTVLSPTSYDLSYRIGD-PIPAAGLDP-----KLMTCDSSDRPR 204
QY 253 -YDMAAVDQ-EF-----TSMVYNTSNVTKLPKPKDMI-GRIDAIDISTGRTLM- 298
DB 205 VYTVTADEYQFSSQLIPSGVKTTLFTANIDALTSVGGELIFSGVTHSIEVDVTVIF 264
QY 299 -----SVERAANYSPVLSGGLVFN--GCTDRYFRALSOETGETLMOTRLATVAS 348
DB 265 IGFDTGEVTVKAVATDFGLTGTNNLVPFLGGTSEITQPIISMKELEVTVYKRGGT--A 322
QY 349 GOAISYEVDGMQYVAIAGG 367
DB 323 GDPISWTVSGTLAVTVIGG 341

RESULT 6
US-09-031-655-10
Sequence 10, Application US/09031655
Patent No. 601759
GENERAL INFORMATION:
APPLICANT: VAKHARIA, VIKRAM
APPLICANT: SNYDER, DAVID B
TITLE OF INVENTION: MENDEL-WHERSAT, STEPHANIE A
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
TITLE OF INVENTION: THEREON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031.655
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/219,262
FILING DATE: 29-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2747-047-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Infectious bursal disease virus
SPRAIN: OH
US-09-031-655-10

Query Match 5.2%; Score 105; DB 3; Length 1012;
Best Local Similarity 22.2%; Pred. No. 0.12;
Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 71 VTNLVHGSTAV-----GPAS-----ETQGTGGTLYGTNTRF 104
DB 1 MTNLMHTQOIVPFRSLMPTGPAISIPDDTLEKHLRSETSYNLTVGDTGSGLIVF 60
QY 105 AVRPDGEIVMRHQTLPD---NMDOECTFEMMT-----NDVQFSTE 145
DB 61 PGFP--GSVGAHHTLOSNGSYQFDQ-----MLTQNLPSVSYNCRVSRSLTVRSSTL 113
QY 146 MEGLOSINP--NAATGERRVLTGVPCKTGMQFDETFELMARDTNYQNMIESI---- 199
DB 114 PGVYALNGTINAVT-----FGSLSEL---TDYSNGMSATANIN 152
QY 200 DENGIVYVNEALIKELDEYDVCPTFLGGRDPSAALNDPSGIYFIPLNNVC----- 252
DB 153 DKIGNVTVGEGVTVLSPTSYDLSYRIGD-PIPAAGLDP-----KLMTCDSSDRPR 204
QY 253 -YDMAAVDQ-EF-----TSMVYNTSNVTKLPKPKDMI-GRIDAIDISTGRTLM- 298
DB 205 VYTVTADEYQFSSQLIPSGVKTTLFTANIDALTSVGGELIFSGVTHSIEVDVTVIF 264
QY 299 -----SVERAANYSPVLSGGLVFN--GCTDRYFRALSOETGETLMOTRLATVAS 348
DB 265 IGFDTGEVTVKAVATDFGLTGTNNLVPFLGGTSEITQPIISMKELEVTVYKRGGT--A 322
QY 349 GOAISYEVDGMQYVAIAGG 367
DB 323 GDPISWTVSGTLAVTVIGG 341

RESULT 7
PCT-US93-10541-2
Sequence 2, Application PC/TUS9310541
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: TRANSGLUTAMINASE GENE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
STREET: 1201 Elm Street, Suite 4500
CITY: Dallas
STATE: TX
COUNTRY: USA
ZIP: 75270-2197

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10541
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126119
FILING DATE: 23-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/969702
FILING DATE: 03-NOV-1992


```

? COMPUTER: IBM AT Compatible
? OPERATING SYSTEM: Windows NT
? SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/356.818A
? FILING DATE: July 19, 1999
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME:
? REGISTRATION NUMBER:
? REFERENCE/DOCKET NUMBER: 2451.002
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (405)652-7252
? TELEFAX:
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 512
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-09-356-818A-2

```

Query Match 4.7%; Score 94.5; DB 4; Length 512;
Best Local Similarity 21.4%; Pred. No. 0.46;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19

```

QY 23 SGDSMTGELMWNFIIPAGEEDM-----GNDYERMTGAMGIIYDPTNLVH 76
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 17 NGRDHHTA-DLCREKLYVRKQ---PRVLLTHEGRNTEA-----SVDSLFRSVLT 63
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 77 YGSTANGPASEFQRTGGTLYGINTPRAPVPDTGELIWRHQTLPBNDMDCEPTEMYVT 136
    ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 64 -----GPASOEA-----GTRKAPPLRDAAVEEDMTATAYVDD-----QCSTLSLQT 109
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 137 NVDVPESTMEGIGIOSINPNATGER-----VLTGPPCKTGIMM--QPAETGEE-L 185
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 106 T-----PANNPIDLYLALSLASTGYGSSPVLGHRILFLPNAWCPADAVYLDSEEEQOEYVL 167
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 186 WANDTYQ---NMISIDENGIVTNEDALIKELDEVDYCPFTL--GGRDWPSAA----- 236
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 162 TGGFTYQGSAAFININPW--FGQFEDGILDIJCLILLIDVNPFLKINAGRDCSRSSPVY 219
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 237 -----LNPDSGIFIFPLNNVCYDMAAYVOEFTSMDY---XNTSNVYIKLRPGKMI 283
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 220 VGRVYSGMVNCDMDGGVLLGRWDMNNYEDGVSPMSYGSVDILRRKNHGCORVYKIGQCVV 279
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 284 GRDIDALDISTGTTLSVERAALANTSPVLTSGGCVLENGCTDRYFR---ALSQETGETIIV 339
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 280 ---FAAVACTVLCGLGIFPRVYVNTNSAHONSMLLT-----EYFNNEGELIQQGRSEKIV 332
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 10
US-08-247-902A-2
; Sequence 2, Application US/08247902A
; Patent No. 5726051

GENERAL INFORMATION:
APPLICANT: Fraij, Bassam M
APPLICANT: Bircichliher, Paul J
APPLICANT: Patterson Jr., Manford K
APPLICANT: Gonzales, Robert A
TITLE OF INVENTION: TRANSGLUTAMINASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: STIDLEY & AUSTIN
STREET: 1201 Elm Street, Suite 4500
CITY: Dallas
STATE: TX
COUNTRY: USA
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,902A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 969702
FILING DATE: 03-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 126119
FILING DATE: 23-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Harre, John A.
REGISTRATION NUMBER: 37345
REFERENCE/DOCKET NUMBER: OMRP B-33590C1PCIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-247-902A-2

```

MOLECULE TYPE: protein
US-08-247-902A-2

Query Match	4.68;	Score 93.5;	DB 1;	Length 548;
Best Local Similarity	21.48;	Pred. No. 0.65;		
Matches 77;	Conservative 48;	Mismatches 148;	Indels 87;	Gaps 19

```

QY 23 SGHDSATGSELBRANFETIRAGEEDDTM-----GNDYERAMTGMQIITYDPTNLVH 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 NGDHDHTA-DLCKRELVYRQ-----PWLTLHREGRNYA-----SYDSLFEVYT 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 YGTAAGPASEFQSTPGCTLYGNTFRFAVRPDTGEIYWRHQTLRPDMNDQCEPMVT 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 -----GPRSOEA-----GTRAFPLRDVAEGDMTATVVD-----QOCTLSLDT 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 NNDVOPSTMEGLQSIINNAATGER-----VLTVPKCTGTMM-OEPAETGEF-L 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 T-----PANAPICLYRLSEASTGYQSSFVLGHFILLFNAKCPAAVYLDSEEGEYVL 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 MARDNYQ-----MLESIDENGIVTNDAILKELDVEYVCETFL--GGRWPSAA---- 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 TQGGELYOGSAKFINIPWN--FGQFEDGJLIDCLILLDVNEPKFNKAGRCSSRSFVY 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 -----LNPPSGIYFPLNNVCYDMAVAQDETSMDY---YNTSNYTKLPPGMDI 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 VGRVYSGMVCNDDQGYLLGMDWNNIGDGVSPMSIGSYDILLRRKNHGCORVRYGGCW 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 284 GRIDAIDISTGRTLMSVERAANSPVLSTGGCVLENGSTDRFR---ALSOETGETM 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 --FAAVACTVLCRGIFPRVYTNNSAHQDNSULL-----EYFNEGEGTQGSKEIM 332

```

RESULT 11
US-08-971-782-4

Sequence 4, Application US/089771782
Patent No. 6043071

GENERAL INFORMATION:

APPLICANT: Wallis, Nicola G.
APPLICANT: Shilling, Lisa K.
APPLICANT: Jaworski, Deborah D.
APPLICANT: Wang, Min
APPLICANT: Mooney, Jeffrey L.
APPLICANT: Debnick, Christine M.
APPLICANT: Zhong, Yi Yi

TITLE OF INVENTION: No. 6043071el G1mt

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Dechert Price & Rhoads

```

STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,782
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,996
FILING DATE: 26-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: GM10024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-971-782-4

Query Match          4.6%; Score 93; DB 3; Length 380;
Best Local Similarity 26.3%; Pred. No. 0.41;
Matches 46; Conservative 26; Mismatches 69; Indels 34; Gaps 9;

QY 66 ITDPVNTLV--HYGSTAVGPASSETGRTGGLTGYTNRFAVRPPDTGEIWMHQTLPRD 123
DB 27 ITGSLKLNLDIFHNKKNVATILTAETDNPFG--YGR-----IVRNDNAEVL--RMVEQKD 78
QY 124 NMDQCEFTPEMNTVNDVQPSSTEM-EGLOSINPNATGERRV--LTGVPCKTG-----TM 174
DB 79 ATDFEQIKKEINTGYVEDNERLFEALKNINTNNAQGEYIITDVIGIFRETGEKVGAYTL 138
QY 175 WQFPAE-----TGEFLWADRTNYQNMIESIDENGIVTNEADAILKEIDVE 219
DB 139 KDFDESLGVNDVALATAESVMRRIRNHKHMV-----NGSVFNPKATYIDIDVE 188

RESULT 12
US-09-309-026-4
; Sequence 4, Application US/09309026
; Patent No. 6204042
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Wang, Min
; APPLICANT: Mooney, Jeffrey L.
; APPLICANT: Debouck, Christine M.
; APPLICANT: Zhong, Yi Yi
; TITLE OF INVENTION: No. 6204042el G1mu
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/309,026
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/971,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: GM10024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-309-026-4

Query Match          4.6%; Score 93; DB 4; Length 380;
Best Local Similarity 26.3%; Pred. No. 0.41;
Matches 46; Conservative 26; Mismatches 69; Indels 34; Gaps 9;

QY 66 ITDPVNTLV--HYGSTAVGPASSETGRTGGLTGYTNRFAVRPPDTGEIWMHQTLPRD 123
DB 27 ITGSLKLNLDIFHNKKNVATILTAETDNPFG--YGR-----IVRNDNAEVL--RMVEQKD 78
QY 124 NMDQCEFTPEMNTVNDVQPSSTEM-EGLOSINPNATGERRV--LTGVPCKTG-----TM 174
DB 79 ATDFEQIKKEINTGYVEDNERLFEALKNINTNNAQGEYIITDVIGIFRETGEKVGAYTL 138
QY 175 WQFPAE-----TGEFLWADRTNYQNMIESIDENGIVTNEADAILKEIDVE 219
DB 139 KDFDESLGVNDVALATAESVMRRIRNHKHMV-----NGSVFNPKATYIDIDVE 188

RESULT 13
US-08-971-782-2
; Sequence 2, Application US/08971782
; Patent No. 6043071
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Wang, Min
; APPLICANT: Mooney, Jeffrey L.
; APPLICANT: Debouck, Christine M.
; APPLICANT: Zhong, Yi Yi
; TITLE OF INVENTION: No. 6043071el G1mu
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,782

```

```

?      FILING DATE:
?      CLASSIFICATION: 435
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: 60/050,996
?      FILING DATE: 26-JUN-1997
?      ATTORNEY/AGENT INFORMATION:
?      NAME: DICKINSON, TODD Q
?      REGISTRATION NUMBER: 28,354
?      REFERENCE/DOCKET NUMBER: G410024
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 215-994-2252
?      TELEFAX: 215-994-2222
?      TELEX:
?      INFORMATION FOR SEQ ID NO.: 2:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 459 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      IS-08-971-782-2

```

	Query Match	4.69;	Score 93;	DB 3;	Length 459;
	Best Local Similarity	26.38;	Pred. No. 0.57;		
	Matches	46;	Conservative	26;	Mismatches 69; Indels 34; Gaps
OY	ITYPDVTLN-V--HGSTAVGPASEFORCTPGCLTGTNTREFAVRPRDGEIYWRHQILPRD	122			
Db	ITGESLKNLIDIFHNKKVATITLAIEDNDPFG---YGR-----TYNDNAEVL--RMVEQKD	157			
	: :		:		
OY	NWDOECFTEMAVTNVDPQSTEM-EGLOSINPNMATGERRV--LTGVPCKTG-----TM	174			
Db	ATDEKQIKIEINTGYVDNERLEAEALKNITNNAGSEYYITDIIGIFRFTGEKVAYTL	217			
	: : : : : : : :		: : :		
OY	WOQFE-----GGEPLMADTVYNQNNIIESIDENGIVTVNEDAILKEELYE	219			
Db	KDFBESLCGVDRVALAATRESVRRIRINKHVV-----NGVSVPNKATYIIDIDE	267			

RESULT 14
US-09-309-026-2
; Sequence 2, Application US/09309026
; Patent No. 6204042
; GENERAL INFORMATION:
; APPLICANT: Wallits, Nicola G.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Wang, Min
; APPLICANT: Mooney, Jeffrey L.
; APPLICANT: Debouck, Christine M.
; TITLE OF INVENTION: No. 6204042el G1md
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastISO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/309,026
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,782
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: GML002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

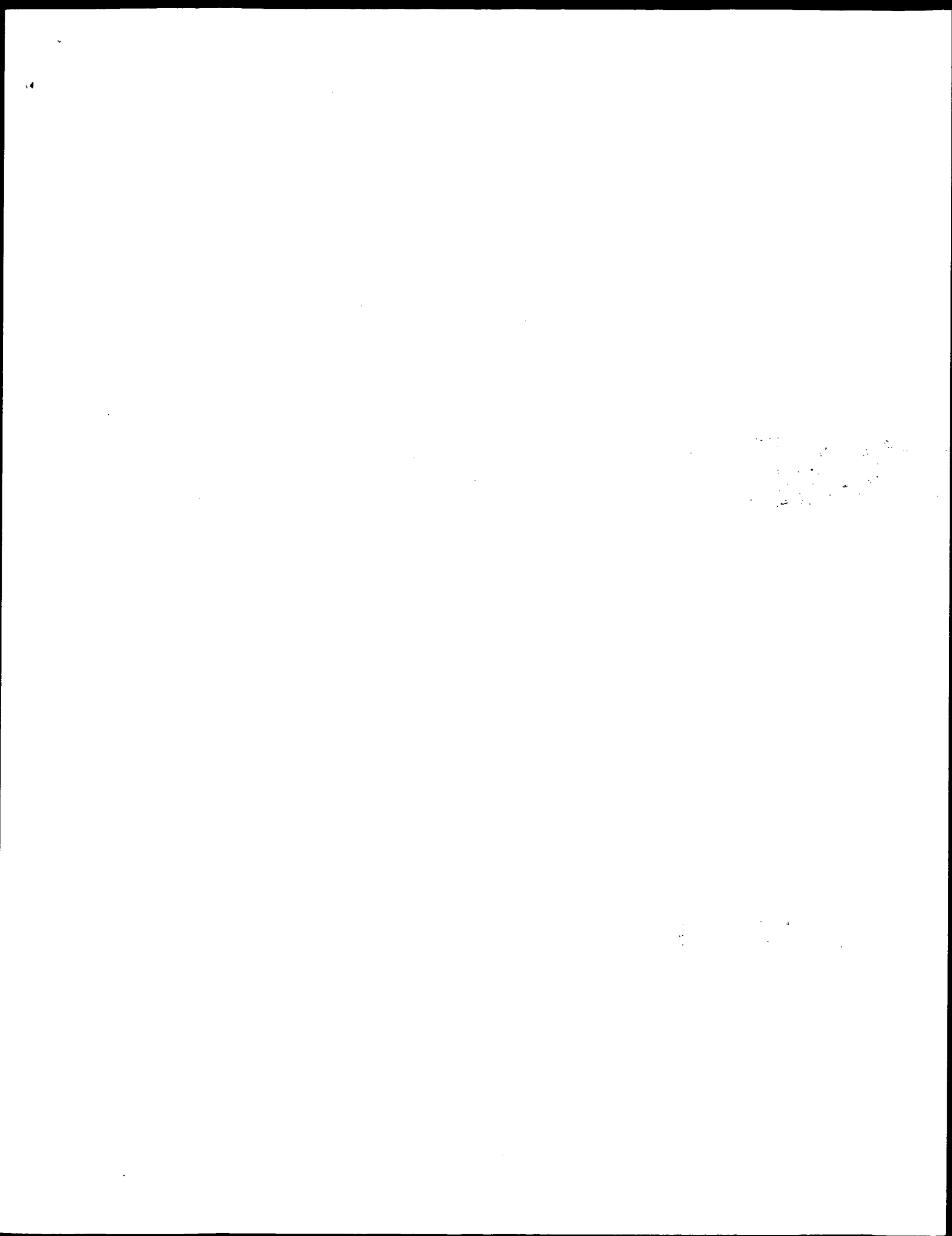
Query Match	4.6%	Score 93	DB 4	Length 459
Best Local Similarity	26.3%	Pred. No. 0.55		
Matches	46	Conservative	26	Mismatches 69
			Indels 34	Gaps 9

QY	66	ITYPDPTNLV--HYGSTAVGPASETORTPGGLTGLTTRFAVPRDGEIYMRHQLPRD	123
	: :	: :	:
Db	106	ITGSLKLNLDIFHNKKNVATLLTAETDNPGE--YGR----IYANDNAEVL--RMEQKD	157
QY	124	NMQDCTFEEMAMTVNDVQPTSEM-BGLDSINPNAATGERRV--LTGVPCKTG-----TW	174
	: :	: :	:
Db	158	ATPEFOKIEKINGTGVFENERLFEALKNINNTNNAOGGYITDVGIFRETGEKVGAYTL	217
QY	175	WQPDAE-----TGEFLMARPTNQNMISSIDEQGIYTVEDALIKELDVE	219
	: :	: :	:
Db	218	KQPDDESGVNDKRALATASVNRKRRIKNNH-----NGSEFVPRKATYIDIVE	267

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:05 ; Search time 56.04 Seconds
(without alignments)
646.425 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556

Perfect score: 2032

Sequence: 1 IVANGYVAGSTQCSYSPFCG.....GMQYVALAGGVSYSGSLNS 377

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	374	18.4	601	2	E95863
2	355.5	17.5	708	2	E52317
3	351.5	17.3	626	2	U00706
4	338	16.6	742	2	U03026
5	338	16.6	742	2	A49340
6	335.5	16.5	573	2	S68591
7	332.5	16.4	738	2	S14270
8	308.5	15.2	623	2	B83399
9	253	12.5	801	1	S00943
10	231.5	11.4	796	1	U01017
11	231.5	11.4	796	2	H90644
12	231.5	11.4	796	2	H85495
13	222.5	10.9	796	2	AC0523
14	209.5	10.3	221	2	A41378
15	207	10.2	809	2	B98314
16	207	10.2	809	2	A12968
17	204.5	10.1	809	2	A55547
18	200.5	9.9	808	1	QPKEX
19	174	8.6	803	2	F83360
20	161.5	7.9	639	2	JC4881
21	148.5	7.3	778	2	G98221
22	148.5	7.3	778	2	A13064
23	130.5	6.4	524	2	A82580
24	122.5	6.0	668	2	C75264
25	119	5.9	407	2	H69064
26	118.5	5.8	386	2	A82284
27	118.5	5.8	799	2	T48889
28	118	5.8	827	2	F64512
29	117.5	5.8	2535	2	AC0304

Query Match	Score	Length	DB ID	Description
1	374	601	2	E95863
2	355.5	708	2	E52317
3	351.5	626	2	U00706
4	338	742	2	U03026
5	338	742	2	A49340
6	335.5	573	2	S68591
7	332.5	738	2	S14270
8	308.5	623	2	B83399
9	253	801	1	S00943
10	231.5	796	1	U01017
11	231.5	796	2	H90644
12	231.5	796	2	H85495
13	222.5	796	2	AC0523
14	209.5	221	2	A41378
15	207	809	2	B98314
16	207	809	2	A12968
17	204.5	809	2	A55547
18	200.5	808	1	QPKEX
19	174	803	2	F83360
20	161.5	639	2	JC4881
21	148.5	778	2	G98221
22	148.5	778	2	A13064
23	130.5	524	2	A82580
24	122.5	668	2	C75264
25	119	407	2	H69064
26	118.5	386	2	A82284
27	118.5	799	2	T48889
28	118	827	2	F64512
29	117.5	2535	2	AC0304

ALIGNMENTS

Query Match	Score	Length	DB ID	Description
1	374	601	2	E95863
2	355.5	708	2	E52317
3	351.5	626	2	U00706
4	338	742	2	U03026
5	338	742	2	A49340
6	335.5	573	2	S68591
7	332.5	738	2	S14270
8	308.5	623	2	B83399
9	253	801	1	S00943
10	231.5	796	1	U01017
11	231.5	796	2	H90644
12	231.5	796	2	H85495
13	222.5	796	2	AC0523
14	209.5	221	2	A41378
15	207	809	2	B98314
16	207	809	2	A12968
17	204.5	809	2	A55547
18	200.5	808	1	QPKEX
19	174	803	2	F83360
20	161.5	639	2	JC4881
21	148.5	778	2	G98221
22	148.5	778	2	A13064
23	130.5	524	2	A82580
24	122.5	668	2	C75264
25	119	407	2	H69064
26	118.5	386	2	A82284
27	118.5	799	2	T48889
28	118	827	2	F64512
29	117.5	2535	2	AC0304

Db 526 RDSPTGDLIMKFKIPSGAIGYPMYTHKGTQYVAL-----YGVG 565

RESULT 4

US0326 alcohol dehydrogenase (EC 1.1.1.-) 72K chain precursor - Acetobacter acetii

C:Species: Acetobacter acetii

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Jun-2000

C:Accession: J50326

R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.

J. Bacteriol. 171, 3115-3122, 1989

A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase sub

A:Reference number: J50326; MUID:89255070

A:Accession: J50326

A:Molecule type: DNA

A:Residues: 1-742 <IND>

A:Cross-references: GB:D50004; GB:M26951; NID:g216193; PIDN:BAAL4058.1; PID:g216194

A:Experimental source: strain K6033

A:Note: amino terminal of mature protein is confirmed

C:Genetics:

A:Gene: adh1

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-742/Product: alcohol dehydrogenase 72K dehydrogenase chain #status predicted <ADC>

Query Match 16.6%; Score 338; DB 2; Length 742;
Best Local Similarity 26.2%; Pred. No. 5e-18;
Matches 114; Conservative 67; Mismatches 152; Indels 102; Gaps 18;

2 VANGVIVAGSTGQYSPFGC--FVSGHDSATGEELMRNYFIPRAGEGD----- 47

203 VAKGLVIGN--GGEFARGFVSAPDAETGKLRFTYTPNNKNEPDHADVADNVLMSKA 260

48 -ETWGNDEYFARMWT-----GANGQITVPYTNLVHGSTAVGPASETORGTGCTLY--- 98

261 YKTWGP--KGAWVRGGGGGTWDSLVYDVPVSDLY--LAVG-----NGSPWNYKRSR 309

99 --GTN-----TRFVRPTGELVWRHQTLPDNDQECTEFEMMTNTVDYOPSTEMEGLQSI 152

310 GIGSNLEPLSGIVALKPTEGEYWHFQATPMDQDYSVOQITLDMFV----- 357

153 NPNATGE--RRVLGVPCKTGTMOQDAETGEFLMARDTNYQNMISIDE--NGIVTNEED 210

359 -----GEMRHVYHAP--KNGFVYLDKATGEFLGKNNYVQNMANGDLPLGRPMYND 411

203 GIYTVNEDAILKELDYEVYVCTFLGGRDPSALNPDSGIYFIFLNNVCY----- 253

412 GIYTLNG-----KFWYGI--PGPLGHNFMAMAYSPKTHLYLPAHOIIPGYKNQVGF 463

254 ----DMMAVDOFTSMQDVNTSNVTKLPKGMIGRIDAIDISTGRITLMSVRAAANYSP 309

464 KPHADSWNVGLDMTKNGLDLPE--ARTATIKDLHGLWLANDPVKMETYWKIDHKGRPMNG 522

310 VLTSGGVFLNGSTDRYFRALSOETGETLMQTRLATVASGAQISYEVGMQYVA----- 363

523 ILATGDLLEFGLANGEFHAYDATNGSDLYKFDQAQSGIILAPMTYSVNGKYVAVEVGMG 582

364 ----TAGGVSYGSG 374

583 GIYPISMGVGGRNSG 597

RESULT 5

A49340 alcohol dehydrogenase (EC 1.1.1.-) precursor - Acetobacter pasteurianus (strain NC11380)

C:Species: Acetobacter pasteurianus

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000

C:Accession: A49340

R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.

J. Bacteriol. 175, 6857-6866, 1993

A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu

A:Reference number: A49340; MUID:94042848

A:Accession: A49340

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-742 <TAK>

A:Cross-references: GB:D13893; NID:g517067; PIDN:BAAL0252.1; PID:g452586

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 16.6%; Score 338; DB 2; Length 742;
Best Local Similarity 26.7%; Pred. No. 5e-18;
Matches 114; Conservative 62; Mismatches 165; Indels 86; Gaps 16;

2 VANGVIVAGSTGQYSPFGC--FVSGHDSATGEELMRNYFIPRAGEGD----- 47

203 VAKGLVIGN--GGEFARGFVSAPDAETGKLRFTYTPNNKNEPDHADVADNVLMSKA 260

48 -ETWGNDEYFARMWT-----GANGQITVPYTNLVHGSTAVGPASETORGTGCTLY--- 98

261 YKTWGP--KGAWVRGGGGGTWDSLVYDVPVSDLY--LAVG-----NGSPWNYKRSR 309

99 --GTN-----TRFVRPTGELVWRHQTLPDNDQECTEFEMMTNTVDYOPSTEMEGLQSI 152

310 GIGSNLEPLSGIVALKPTEGEYWHFQATPMDQDYSVOQITLDMFV----- 357

153 NPNATGE--RRVLGVPCKTGTMOQDAETGEFLMARDTNYQNMISIDE--NGIVTNEED 210

358 -----NEMRHVYHAP--KNGFVYLDKATGEFLGKNNYVQNMANGDLPLGRPMYND 411

211 AILKELDYEVYVCTFLGGRDPSALNPDSGIYFIFLNNVCY-----DMAA 257

412 GIYTLTKFVWYGI--PGPLGHNFMAMAYSPKTHLYLPAHOIIPGYKNQVGF 463

258 VDOFTSMQDVNTSNVTKLPKGMIGRIDAIDISTGRITLMSVRAAANYSPYLTSGGV 317

472 VGLDMTKNGLDLPE--ARTATIKDLHGLWLANDPVKMETYWKIDHKGRPMNGVLATGDL 530

318 LFNSTDRYFRALSOETGETLMQTRLATVASGAQISYEVGMQYVA-----TAGG 367

531 LFOGLANGEFHAYDATNGSDLYKFDQAQSGIILAPMTYSVNGKYVAVEVGMGIPISMG 590

368 GVSYGSG 374

591 GVGRTSG 597

RESULT 6

S68591 methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylobacillus methylotrophus (str

C:Species: Methylobacillus methylotrophus

A:Variety: strain W3A1

C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999

C:Accession: S68591

R:Xia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.

J. Mol. Biol. 259, 480-501, 1996

A:Title: Determination of the gene sequence and the three-dimensional structure at 2.

A:Reference number: S68591; MUID:96256524

A:Accession: S68591

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-573 <XIA>

A:Cross-references: EMBL:U41040; NID:g1127819; PIDN:AAA83765.1; PID:g1127820

A:Note: the authors did not translate the codons for residues 1-2

C:Keywords: oxidoreductase

Query Match 16.5%; Score 335.5; DB 2; Length 573;

Best Local Similarity 24.3%; Pred. No. 5.4e-18;

Matches 108; Conservative 65; Mismatches 166; Indels 101; Gaps 16;

3 ANGVIVA-----GSTCQYSPF-----GC-----FVSGHD 26

```

Db 125 ANGHLALDLAKTKINMEVEYCDKRVSTLTQAEPFAKDTVLKMGSCGALGVGAVNAFD 184
QY 27 SATGEELMR-----NYFIPRAGE--EGDETWGNDYEAARMATGA---WGQ 65
Db 185 LKTELKMRFAATGSDSVRLADFNANPHYCGFGIGTWTGCD---AMKIGGGTMMGW 241
QY 66 ITYPPTNLVHYGSTAVGAPSETOGTPGG---TLYGNTFRFAPVPTDGEIWRHQT 120
Db 242 YADPKNLTFYGGSGNAPANNETWR--PGDNKMTMTITWGRDL-----DTGMKMGYOKT 293
QY 121 PRDNWDQCEFFEMVTVNDVQPTSTEMEGLOSINPNMATGERRVLTGVPCKTGTMMQPDAL 180
Db 294 PHDEMWDAGYVQWVMDL-----QPVN-----GAMPPLSHIDNGILLYTLNRE 336
QY 181 TGEFLMARDTN-YONMIESIDENGIYVNEADALLKELDVE-YDVCPTFLGGRDMPAALN 238
Db 337 NGNLIVAEKVDPANVNAFKKDLKTGTPVRPEFATRMDHGNICISAMFHNQVSDYD 396
QY 239 PDGSIYIPLNNVCYDMAVDOFTSMQVNTSNTKLP---PGKDMIGRIDALIDISTG 294
Db 397 PESTTLVAGNLHICMDPEPMLPYRAGQFVGAATLAMPGPNGPFTKEMQIRAFDLTTG 456
QY 295 RTLMSVERAANYSPLYSTGGVLFNGGDRFRALSOETGELTMQTRLATVASGAISY 354
Db 457 KAKTKKKEKFRANGGILTYTGGVLWATATLDGYLKALDNKDKELMFKMPSGGIGSPMTY 516
QY 355 EVDGMQVVAIAGGVSYSG 374
Db 517 SFKGKQYI-----GSMGVG 531

```

RESULT 7

```

S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter pol
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S14270
R:Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiyama
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane-bc
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BA00528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72k and 44k chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-33/Domain: signal sequence #status predicted <SIG>
F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAM>

```

```

Query Match 16.4%; Score 332.5; DB 2; Length 738;
Best Local Similarity 26.8%; Pred. No. 1.3e-17;
Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;

```

```

QY 2 VANGVIVAGSTCOYSPFGC--FVSGHDSATGELMNRNYFPRAGEGD-----47
Db 204 IAKGRVITIGN--GSEFGARGVSAFADATGKVDNRFFVTPPKKDEPPDASVLMKAY 261
QY 48 ETWGNDEARMMTG-----WGQITYPVNLVHYGSTAVGAPSETOGTPGGTL 97
Db 262 QFMS-----PLGAWTRQGGGGTWDLSIVDPADVLYLGAGNSPMNKKRSEKGN 314
QY 98 YGNTFRFAPVPTDGEIWRHQTLPDRNDQCEFFEMVTVNDVQPTSTEMEGLOSINPAA 157
Db 315 LFLGSIYALKPTEGVEWHFQETPMQDMFTSDQIMTLDLP-----357
QY 158 TGERRVLTGVPCKTGTMMQFDATGELMARDTNQNMIESID-ENGIVTVNEDAILKEL 216

```

```

Db 358 NGETRHVIVHARKNGFEYIIDAKTGEFISGKNVYVYVNNASGLDPKTRPIYNPDALYTLT 417
QY 217 DVEYDVCPTFLGGRDMPAALNPDGSIYIPLNNVCYDMAVDOFT-SMDVYNTS---N 272
Db 418 GKEWIGIPBDLGGHFAAMAFSPKGTGLYIIPQOVPFLTYNOVGFTTHPHPSMNLGLDMN 477
QY 273 VTKLPG-----KDMIGRIDALIDISTGRTLMSVERAANYSPLYSTGGVLFNGCTD 324
Db 478 KVGIDPSPEAKQAFVKDLKGMIVAMDPOKQAEAMRVHDKGPMNGIILATGDDLLFQGLAN 537
QY 325 KYFRALSOETGELTMQTRLATVASGAISYEVDGMQVVAIAG 367
Db 538 GEFHAYDATNGSDLFHFAADSGIAPPVYTLANGKQYAVEVG 580

```

RESULT 8

```

B83399
quinoxaline alcohol dehydrogenase PA1982 [Imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83399
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: B83399
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-623 <STO>
A:Cross-references: GB:AE004624; GB:AE004091; NID:9947973; PIDN:AA05370.1; GSPDB:GN
A:Experimental source: strain PA01
A:Genetics:
A:Gene: exaB; PA1982

```

```

Query Match 15.2%; Score 308.5; DB 2; Length 623;
Best Local Similarity 22.9%; Pred. No. 7.3e-16;
Matches 88; Conservative 68; Mismatches 175; Indels 53; Gaps 10;

```

```

QY 26 DSATGEELMRNYFIPR-----AGEGDETWGNDYEA-----KMTGA---WG 64
Db 224 DPDTGEELIMRPFVEGHHMGRNGKDSITVGVKAPSWPDDNSPTGVESSHGAPWQ 283
QY 65 QITYPVNLVHYGSTAVGAPSETOGTPGG-----TLGNTFRFAPVPTDGEIWRHQ 118
Db 284 SASPEAETNTIIVGAGNPGPNNTAKGAGNPHDYSLY-TSGQGVDPSSGGEVKAFTQ 342
QY 119 TLPDRNDQCEFFEMVTVNDVQPTSTEMEGLOSINPN-----AATGERRVLTGVPCKTGT 173
Db 343 HTPNDAMQFSGNNELVLEFDYKAKGKIKATAHADNRGFFVYVDRSGKLGQNAFFVDNI 402
QY 174 MW--QFDEATGEFLMARDTNQNMIESIDENGIYTVNEDAILKELDVEYVCPFLGGRD 231
Db 403 TWASHIIDKLTG---RVERREGORPLPEPG-----OKHGRAVEYSPFLGKGN 447
QY 232 WPSAALNPDGSIYIPLNNVCYDMAVDOFTSMQVNTSNTKLPKPGKMIGRIDALID 291
Db 448 WNPAAVYSDTDLTFYVPAHMKEDYVTEVSTYTKSAYLGAGFRKRRKRYDDHVSLSRAMDP 507
QY 292 STGRTLMSVERAANYSPLYSTGGVLFNGGDRFRALSOETGELTMQTRLATVASGA 351
Db 508 VSGKVVHEHKKHPLMLMAGVILATAGNLVFTGQDYFAFAFAKSGKELMKQTQSSGIYSP 567
QY 352 ISYEVDGMQVVAIAGGVSYSGSL 375
Db 568 ITWBQDGEQYLGVP---VGYGAV 588

```

RESULT 9

```

S00943
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calco

```


[illegible][illegible]

```

Db      559  CRYMFFHQMKREGIETPPSEQGLTVFPNGNLGMEEMOGISVDPNREVAINPMALPEVSKLI 618
QY      240  -----DSGI-----YFPLNNVCYDMAVDOEFTSDMYNTSNV 273
Db      619  PRGPGNPMBPQKADKAGTGESGIGPOGYPVYVLTN-----PFLS-----P 659
QY      274  TKLPFGKMDIGRIDADISTGRTLSVERAANYS-----PV-----LSTG 314
Db      660  FELPKQKAPMGWISLADLKTNEVWKKRIGTPQDSMPPEMPVPYFPNNGMPLGPISTA 719
QY      315  GGVLEVNGGT-DRYFRLASQETGETLMQTRLATVASGQA--ISYEVDGQMYAI--AGGSVS 370
Db      720  GNVLEFIATADNYLRAVYMNNGEIKMGRLP--AGQAATPTMYEVNGQYVVISAGHGS 777
QY      371  YGS 373
Db      778  FGT 780

```

RESULT 12
H85495
glucose dehydrogenase [imported] - *Escherichia coli* (strain O157:H7, substrain EDL9333)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 14-Sep-2001
C:Accession: H85495
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Hiller, L.; Griebcke, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoustis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85495
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <STO>
A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AA634428.1; GSPDB:GND0145; UMGF:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics
A:Gene: gdcJ
A:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

	Query Match	Similarity	11.4%	Score	231.5	DB	2,	length	796;
	Best Local	Similarity	23.4%	Pred.	No. 9e-10				
	Matches	Conservative	53;	Mismatches	154;	Indels	163;	Gaps	23;
QY	6	VIVAGS-TCQYS-PFGCFVSGHDSANGELMRNYFLTPRA-----GEEGETWGNDE	55						
	:	:::::	:	:	:	:	:	:	:
Db	346	IYAVGSVDNFTSTRETSYIRGFIDVTNGELLMA--FDPGAKDNPAISPDHTFFMS---	400						
QY	56	ARMNTCANCOILTYDDVTNLVNHGSIANGPASETGRGPGLYLGTINTRAVERDPTGEIYW	115						
	:	::	:	:	:	:	:	:	:
Db	401	---PNSMAPAAYADKDLVLLPMGVTTPLDMGNNRPEDERYASSI-LALNATGGKLAW	455						
QY	116	RHOITLPRMDDECETEMNAVTVNVQOSTEAMEGLSOSINPAATGERRVLTGVCRTGTM	175						
	:	:::	:	:	:	:	:	:	:
Db	456	SQTVAHHLDMDND-----LPAQPL-----LAQITVA--GQKVPVIYAPAKTGNI	498						
QY	176	QPDATGEFL-----WADTNNTONMIESIDENGI-----VTVNDAI	212						
	:	:	:	:	:	:	:	:	:
Db	499	VLDRRNGELLVPAPEKPYPVOGAAGDVYTPTOPFSELSFRPKDLSDGMWGATMFQDLV	558						
QY	213	LKEVL-DVEYD-----VCPFLGGRMPASALNP-----	239						
	:	:	:	:	:	:	:	:	:
Db	559	CYVMFHQARYBEIFLPSPESOGTLVFPGNLGHFEFGGISVDPNREVALIANPALPVFSKLI	618						
QY	240	-----DSGI-----YFLPLNNVCYDMAAAVDOETSMDVYNTSNV	273						
	:	:::	:	:	:	:	:	:	:
Db	619	PRGPNPMEOPKDAKGTFGESGIQPYQGVYGYVTLN-----PLTS-----P	659						
QY	274	TSLRPGKMIGRIIALIDISTGRITMSVERAANY-----PV-----LSTG	314						
	:	:::::	:	:	:	:	:	:	:
Db	660	FELPKQAPAMGISTLDTKTNEVWKRICTPODSWFPPVPPVFNMGMPIMJGSPISTA	719						

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:50 ; Search time 27.86 Seconds
(without alignments)
523.951 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556

Perfect score: 2032
Sequence: 1 IVANGYIVAGSTGQYSPFGC.....GMOYVIAAGGVSYSGSLNS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	361	17.8	600	1	XOFE_PARDE
2	351.5	17.3	626	1	DMH1_METEX
3	341.5	16.8	626	1	DMH1_METOR
4	340	16.7	631	1	DMH1_PARDE
5	338	16.6	739	1	DHET_ACCEU
6	338	16.6	742	1	DHET_ACCEC
7	332.5	16.4	571	1	DMH1_METME
8	332.5	15.2	738	1	DHET_ACCEP
9	308.5	15.2	623	1	EXAA_PSEAE
10	298.5	14.7	757	1	DHET_GLUOX
11	293	12.5	801	1	DHGA_ACICTA
12	231.5	11.4	796	1	DHG_ECOLI
13	204.5	10.1	809	1	OUJA_ACICTA
14	200.5	9.9	808	1	DHG_GLUOX
15	180.5	8.9	790	1	OUJA_XANCI
16	118.5	5.8	799	1	ASEK_STRO
17	118	5.8	827	1	YK23_METTA
18	115.5	5.7	353	1	YKAL_BACSU
19	115.5	5.7	354	1	CADN_MOUSE
20	115	5.7	443	1	PORD_PSEAE
21	114	5.6	796	1	COPP_SCHPO
22	110.5	5.4	317	1	CADN_RAT
23	108.5	5.3	354	1	CADN_HUMAN
24	105.5	5.2	593	1	SPG2_STRSP
25	102	5.0	790	1	PLMN_PIG
26	102	5.0	1645	1	OMPB_RICVY
27	101	5.0	392	1	YFGL_ECOLI
28	101	5.0	1012	1	POLG_IBDVO
29	99.5	4.9	466	1	MMO8_RAT
30	99.5	4.9	807	1	AFSK_STRGR
31	98.5	4.8	1016	1	PMPH_CHLTR
32	96.5	4.7	465	1	MMO8_MOUSE
33	95.5	4.7	746	1	FEPA_ECOLI

34	95	4.7	570	1	SYG_PYRHO
35	94.5	4.7	746	1	FEPA_PSEAE
36	94	4.6	546	1	CH60_BURCE
37	94	4.6	546	1	CH60_BURVI
38	93.5	4.6	545	1	CH63_BRAJA
39	93.5	4.6	687	1	TGIC_HUMAN
40	93.5	4.6	2248	1	CYAL_DROME
41	93	4.6	537	1	ANPC_BOVIN
42	93	4.6	1365	1	GTPC_STRO
43	92.5	4.6	954	1	FLRY_CAUCR
44	91.5	4.5	760	1	METE_MYCLE
45	91	4.5	504	1	VL2_HPV63

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description
1	XOFE_PARDE	361	17.8	600	1	XOFE_PARDE
2	XOFE_PARDE	351.5	17.3	626	1	DMH1_METEX
3	XOFE_PARDE	341.5	16.8	626	1	DMH1_METOR
4	XOFE_PARDE	340	16.7	631	1	DMH1_PARDE
5	XOFE_PARDE	338	16.6	739	1	DHET_ACCEU
6	XOFE_PARDE	338	16.6	742	1	DHET_ACCEC
7	XOFE_PARDE	332.5	16.4	571	1	DMH1_METME
8	XOFE_PARDE	332.5	15.2	738	1	DHET_ACCEP
9	XOFE_PARDE	308.5	15.2	623	1	EXAA_PSEAE
10	XOFE_PARDE	298.5	14.7	757	1	DHET_GLUOX
11	XOFE_PARDE	293	12.5	801	1	DHGA_ACICTA
12	XOFE_PARDE	231.5	11.4	796	1	DHG_ECOLI
13	XOFE_PARDE	204.5	10.1	809	1	OUJA_ACICTA
14	XOFE_PARDE	200.5	9.9	808	1	DHG_GLUOX
15	XOFE_PARDE	180.5	8.9	790	1	OUJA_XANCI
16	XOFE_PARDE	118.5	5.8	799	1	ASEK_STRO
17	XOFE_PARDE	118	5.8	827	1	YK23_METTA
18	XOFE_PARDE	115.5	5.7	353	1	YKAL_BACSU
19	XOFE_PARDE	115.5	5.7	354	1	CADN_MOUSE
20	XOFE_PARDE	115	5.7	443	1	PORD_PSEAE
21	XOFE_PARDE	114	5.6	796	1	COPP_SCHPO
22	XOFE_PARDE	110.5	5.4	317	1	CADN_RAT
23	XOFE_PARDE	108.5	5.3	354	1	CADN_HUMAN
24	XOFE_PARDE	105.5	5.2	593	1	SPG2_STRSP
25	XOFE_PARDE	102	5.0	790	1	PLMN_PIG
26	XOFE_PARDE	102	5.0	1645	1	OMPB_RICVY
27	XOFE_PARDE	101	5.0	392	1	YFGL_ECOLI
28	XOFE_PARDE	101	5.0	1012	1	POLG_IBDVO
29	XOFE_PARDE	99.5	4.9	466	1	MMO8_RAT
30	XOFE_PARDE	99.5	4.9	807	1	AFSK_STRGR
31	XOFE_PARDE	98.5	4.8	1016	1	PMPH_CHLTR
32	XOFE_PARDE	96.5	4.7	465	1	MMO8_MOUSE
33	XOFE_PARDE	95.5	4.7	746	1	FEPA_ECOLI

OY

2 VANGYIVAGSTGQYSPFGCFSVSGHDSA-----TGEIWMNYI-----39

```

Db 180 VKDVLVIGSGEG-----VRGMMTALNTLDSSEAKWASTGPDEELVDPETTHLTK 234
QY 40 PRAGEEDDETNGNDYERAMMTGA---WGQITYPVNLVHYGSTANGPASETORPGT 96
Db 235 PIGADSSLSNWECD---QMOIGGTTWGFSTYDPLNLVYGGTNGSTWMPQSQR--PGDN 289
QY 97 LYGNTFRFAVRPDTGELVWHRQTLPRDNDMOECTFEKMTN--VDYOPSTMEGLQSI 154
Db 290 KM-SMTIARDADLTGMKMFYQMTPRDEMDYDGVNEMILLNQYVDO-----335
QY 155 NAATGERVLTGVPCKTGTMQDAETGEFLMARDTNYOMIE-----SIDENGIYVNE 209
Db 336 -----ERKLLTHFD-RNGLAITYTDRTEGELVAE--KYRPPVWMTTGVMDPNS-ETYGR 386
QY 210 DAILEKEIDVEXD-----VCPTEFGRRDMPASALNPDGTFEPLNVCYDMAVDOE 261
Db 387 PAVVAESTQNGEDENTGTGVCALAGTQDOOPAFSPKTNLFYVTHVNCMDYEPFVA 446
QY 262 FTSMDVY--NTSNVTKLPCKDMIGRIDAIDISTGRTLSVERAANYSPVLSTGGVLF 319
Db 447 YTAGOPYVGCATLSMYPAPNSHGMNFIAHNTTGEIKMSVPEQPSVWGALATAGDYVF 506
QY 320 NGGTDYFRALSGEFTETIMQTLAVASGQALSYVDQMUYALAG-----GVSYSGS 374
Db 507 YGTEGLTKLPVDAQTESEELKFTKPTSGIIGNWYTEHGKQYVILSGVWGAGIGLAG 566
QY 375 LNS 377
Db 567 LTN 569

RESULT 2
DHML_METEX STANDARD: PRT: 626 AA.
AC P16027;
DT 01-APR-1990 (Rel. 14, Last Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
  alpha subunit) (MDH).
GN MOXF.
OS Methylobacterium extorquens.
OC Bacteria: Proteobacteria; alpha subdivision: Rhizobiales group;
  CC Methylobacterium group; Methylobacterium.
OX NCBI_taxid=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AM1 / NCIB 9133;
RX MEDLINE=90337342; PubMed=2116368;
RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
RT "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
  moxf genes involved in methanol oxidation.";
RL Gene 90:173-176(1990).
RN [2]
RP SEQUENCE OF 28-53.
RC STRAIN-AM1 / NCIB 9133;
RX MEDLINE=89350892; PubMed=2504152;
RA Nunn D.N., Day D., Anthony C.;
RT "The second subunit of methanol dehydrogenase of Methylobacterium
  extorquens AM1.";
RL Biochem. J. 260:857-862(1989).
RN [3]
RP DISULPHIDE BONDS.
RX MEDLINE=95384759; PubMed=7656012;
RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
RT "The active site of methanol dehydrogenase contains a disulphide
  bridge between adjacent cysteine residues.";
RL Nat. Struct. Biol. 1:102-105(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RX MEDLINE=95253818; PubMed=7735834;
RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;

```

```

RT "The refined structure of the quinoprotein methanol dehydrogenase
  from Methylobacterium extorquens at 1.94 A.";
RL Structure 3:177-187(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
  reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
  OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
  MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation-
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M31108; AAA25380.1; -
DR PIR: S07908; J00708.
DR PIR: J00706; J00706.
DR HSSP: P38539; 4AAB.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT CHAIN 1 27
FT DISULFID 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 130 131
FT DISULFID 413 442
FT ACT_SITE 330 330
FT ACT_SITE 330 330
SQ SEQUENCE 626 AA; 68434 MW; 64988DDAFDZAD34C CRC64;

Query Match 17.3%; Score 351.5; DB 1; Length 626;
Best Local Similarity 26.2%; Pred. No. 8.8e-20;
Matches 106; Conservative 69; Mismatches 169; Indels 61; Gaps 14;

QY 1 IVANGIVASTCOYSPGCFVSGHDSATGEELMRNYFI-----PRAGE 44
Db 191 VKDKVILIGSSGAEALGVRG-YLTAVDYKTEQVWRVATGPDKDLLASDFNIKNHYGQ 249
QY 45 E--GDETGWNDYERAMMTGA---WGQITYPVNLVHYGSTANGPASETORPGTLYG 99
Db 250 KGLSTGWEGD---AMKIGGGTNGWYAYDPGTLNLYFGTGNAPWNETMR--PGDNKN- 303
QY 100 TTRFAVRPDTGELVWHRQTLPRDNDMOECTFEKMTN--VDYOPSTMEGLQSI 159
Db 304 TMTIFGRDADLTGKAKRQYKQTPRDEMD-----YAGVAVMLLSOKD-----KKG 347
QY 160 ERRVLGVPCKTGTMQDAETGEFLMAR--DTNYOMIESIDENGIYVEDALIKEL 216
Db 348 KARKLLTHPRNRNGIYVTLRFDGALVANSANKLDFT--VNFKSVDLKTGGPVRDPEGTGA 405
QY 217 D-VEYDVCPEFLGRRDMPASALNPDGSIYFIPLNNCYDMAVDOETFSMDYINNSVTK 275
Db 406 DHIKADICPSAMVYHNGQHDSDPKRELFPMGNHNCIMWEPMLPYRAGQEFFVATLMM 465
QY 276 LPPGK-----DMIGRIDAIDISTGRTLSVERAANYSPVLSTGGVLFNGCTRRYFRA 329
Db 466 YRPKPKDRQNYEGLGQIKAVNATGTGYKKKKMERFRAVWGCTAATGADLVFTGTLGCIYKA 525
QY 330 LSGTEGTELMQTRLATVASGQALSYVDQMUYALAGGVSYSGS 374
Db 526 RDSDTGDLMLKFKIPSGAIGYPMYTHNKGTQYVAI-----YGVGV 565

RESULT 3
DHML_METEX

```

```

ID      DHML METOR          STANDARD;          PRT:    626 AA.
AC      P15279:
DT      01-APR-1990 (Rel. 14, Created)
DT      01-APR-1990 (Rel. 14, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE      alpha subunit) (MEDH).
GN      MOXF.
OS      Methylotacterium organophilum XX.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Methylotacterium group; Methylotacterium.
OX      NCBI_TaxId=410;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RC      STRAIN-ATCC 27886 / DSM 760 / NCIB 11278;
RX      MEDLINE-89008094; PubMed-2459109;
RA      Machlin S.M., Hanson R.S.;
RT      "Nucleotide sequence and transcriptional start site of the
RT      Methylobacterium organophilum XX methanol dehydrogenase structural
RT      gene";
RL      J. Bacteriol. 170:4739-4747(1988).
CC      -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC      reduced acceptor.
CC      -I- COFACTOR: POQ.
CC      -I- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC      OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC      -I- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC      MEMBRANE.
CC      -I- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; M26529; AAA50289.1; -.
DR      HSSP; P38539; AAAH.
DR      InterPro; IPR001472; Bac_POQ.
DR      InterPro; IPR002372; Bac_POQ_repeat.
DR      Pfam; PF01011; Bacterial_POQ_7.
DR      PROSITE; PS00363; BACTERIAL_POQ_1; 1.
DR      PROSITE; PS00364; BACTERIAL_POQ_2; 1.
KM      Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
FT      SIGNAL              1      28
FT      CHAIN               29      626   METHANOL DEHYDROGENASE SUBUNIT 1.
FT      DISULFID            130     131   BY SIMILARITY.
FT      DISULFID            413     442   BY SIMILARITY.
FT      ACT_SITE            330      330   BASE (POTENTIAL).
SQ      SEQUENCE           626 AA; 68677 MW; 8768FB8371E5DF CRC64;

Query Match                16.8%; Score 341.5; DB 1; Length 626;
Best Local Similarity      26.2%; Pred. No. 5,2e-19;
Matches 106; Conservative 67; Mismatches 171; Indels 61; Gaps 14,

OY      1 IVANGVIYAAGTCQISPFCEVSGHDSDATGEELMRYFI-----PAGE 44
DB      191 VVKDKVITIGSSAEELGVNRG-YLLTAADVKTGGWVRAYATGPDKDLLADDFNKNAHYGO 249
OY      45 E--GDETWGNDYEARMNTGA--WGQITVDPTNLVHGYSWAVGPASEFGPGGTILG 99
DB      250 KGLIGRTATMGCD---AKKIGGGTNMGVAYADPGRNLIIFFGTGNPAPNNETMR--PGDNKM- 303
OY      100 TMTREAVRDGTGEIVWRHOTLPEDNMDOECTEMVTNVADVOSTEMEQLQSINPAAATG 159
DB      304 TWMTIFGRADDTGEAKRGYOKTPHDEWDYAGVVWM-----PSEKOD-----KDG 347
OY      160 ERRVVLGVGCCKTGTMNQDPDAEGEFIMAR---DTNQNMMIESIDENGIVTVNEDALIKEL 216
DB      348 KTKKLTLTHPDRIKGIYTTLDRTGDALVSANKLDDT--VNVEFKYIDLTKGTGPVPDPEIGTKM 405

```

```

QY      1 IVANGVIYAGSGTGYSPF-----GC-----FVSGHDSATGTEELRNRYFI---- 39
Db      276 LPPEK-----DMGRIDAIIDISTGRITLMSVERAANYSPLYSTGGGVLFNGGDTRYRA 329
QY      406 DLAHKDVCPSAMGHNHGCHSDYPDKRELFEFGINICHMDPEPMLPYRAGCFPVGATILNM 465
Db      466 YPGKGGRONTEGLGQIKAYNAITGSYKWEKMERPAWVGTLTAGDLVFYGTLDYLKA 525
QY      330 LSGETGETLWTRLATVASGOAISYENDGMOTVYIAGGVSYSGS 374
Db      526 RDSPTGLIMFKRIPSGAIGYPMTYTHKGTOYVAI-----YYGWG 565

RESULT      4
DHMT_PARDE  ID DHMT_PARDE          STANDARD:          PRT:         631 AA.
AC           PI2293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
   alpha subunit) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria: Proteobacteria; alpha subdivision; Rhodobacter group;
CC Paracoccus.
CX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RX MEDLINE=87307969; Pubmed=3114231;
RA Haux N., de Vries G.E., Maurer K., Hoogendijk J., Stouthamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
   structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975(1987).
CC -1 CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
   reduced acceptor.
CC -1 COFACTOR: POO.
CC -1 SUBCELLULAR LOCATION: Periplasmic.
CC -1 MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
   ON METHANOL (IN P.DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
   TOTAL CELL PROTEIN).
CC -1 SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
   or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: M17339; AAA88366.1; -.
DR HSSP: P38539; AAAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductases; POO; Signal; Methanol utilization; Periplasmic.
FT SIGNAL                      1       32
FT CHAIN                       33      631
FT DISULFID                     135     136
FT DIISULFID                    418     447
FT ACET SITE                    335     335
SQ SEQUENCE                   631 AA; 69799 MW; 0934DC93FFC5730B CRC64;

Query Match              16.7%; Score 340; DB 1; Length 631;
Best Local Similarity    25.0%; Pred. No. 6, 9e-19;
Matches 107; Conservative 65; Mismatches 166; Indels 90; Gaps 15

```

```

Db 178 IMENSDIKVSTLTIAPIVKIDLVVSSGSAELGVRGIVAYDVKSGEMRRARATGPDE 237
OY 40 -----PRAGEE--GDETGNDEYEAAMTGA---WGQITYPVTNLVHGSTAV 82
Db 238 ELLIAEDFNAPNPHYGKKNIGLETWED---AMKIGGTMMGWAYDEVDLFFYSSGNP 294
OY 83 GPASEPQRTGPGGLTYTNTREAVRPTGELVWHQTLPRDNDQECTFEKMTVNDVOP 142
Db 295 APNMEYTR--PGDNKW--TMAIWGREATTGEKFAVQKTPHDEM-----YAGVVM 343
OY 143 STEMEGQSINPNATGERRVLTGVPCKTGTMQFQDAETGFEFLARDTNYQNMIESIDE- 201
Db 344 LSEBEDQ-----GQMKRLTHPRNGIVYTLDRKNSDLISA-----DKMDDT 386
OY 202 -NGIVYNEDAILKELDEY-----DVCPTFLGRDMPASALNPDGTYPIPLNV 251
Db 387 VNNVKEVQDLPTGVRDPERGTRMDHKARDICPSAMGYHNGHSDYPERKRVFLGNIHI 446
OY 252 CYDMAVADQFTSMQVNTSNVTKLPGKDM-----IGRIDAIDISTGRTLSYERAAN 306
Db 447 CCMQEPMLPYRAGQFVGATLTMYPGPKATAEKAGQIKAYDAISGEMKMEKMERFSV 506
OY 307 YSPVLSTGGGVLNFGTDRFRLISOETGETIMQTRLATVASGQAISEVDGMQYVIAAG 366
Db 507 WGGTMTAGLTYVTLIDGFIKARDSDPTGDLMLKFLPSGVIGHPMYKHGDRQYVAI-- 564
OY 367 GGVSYSG 374
Db 565 ---MYGVG 569

RESULT 5
DHET_ACCEU STANDARD; PRT; 739 AA.
AC 044002; 007952;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADH.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae.
OX NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=DES11 / DSM 6160;
RA Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC
CC EMBL; X82894; CAA58066.1; -
CC EMBL; Y09480; CAA70688.1; -
CC HSRP; 09Z4J7; 1FLG.
CC InterPro: IPR001479; Bac_POQ.
CC InterPro: IPR002372; Bac_POQ_repeat.

```

```

DR InterPro: IPR00345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_POQ_6.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C_1.
KW oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1
FT CHAIN 36
FT BINDING 651 651
FT BINDING 654 654
FT METAL 655 655
SQ SEQUENCE 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;

Query Match
Best Local Similarity 27.2%; Pred. No. 1,2e-18;
Matches 108; Conservative 54; Mismatches 185; Indels 50; Gaps 10;

OY 2 VANGVIAGSTCOYSPGCG--EYSGHDSANGCELMRYFIIPRAGEGDETWGND----- 53
Db 204 IAKGRVILGN--GGSEFGARGFVYAFDAETGKVDMRFFTAAPNKNPDHPTASDVLMNKA 261
OY 54 YEARNMTGA-----WGQITYPVTNLVHGSTAVGAPASTQGTGRTIGTINTR 103
Db 262 YQWSPFGAMTRQGGGQSTWDSIVYDPAVLVLGVNGSPWNYKRSSEKGNLFLGSI 321
OY 104 FAVRPDTGEIYWRHOTLPNDNDQECTFEKMTVNDVOPSTEMEGLQSNPNATGERRV 163
Db 322 VALKPEYGEVYWHQETPRMQWDETSVQOIMTLDPINGET-----RHV 365
OY 164 LTGVPCCKTGMQFQDAETGFEFLARDTNYQNMIESID-ENGIVYNEDAILKELDEYDV 222
Db 366 IVHAP-KNGFFYIIDAKTGEFISGKNVYVWMAAGLDPKTRPYNDALTYTLGKEMYG 424
OY 223 CPTFLGRDMPASALNPDGTYPIPLNVQCYDMAVADQFT-SMDVYNTS---NVTLLRP 278
Db 425 ITPDGLGHNFAAMAFSKTLVYIIPAOQVFLTYNOVGFTPHPDSSNKLGLDMKVGIPD 484
OY 279 G-----KDMIGRIDAIDISTGRTLSYERAANSPVLSYSGGVLFNGTDRYFRAL 330
Db 485 SPEAKQAFVAKDLGWIYAWMPQKQAEAMRYVDHGMWNGILATGDLFLFGLANGEFHAY 544
OY 331 SQETGETIMQTRLATVASGQAISEVDGMQYVIAAG 367
Db 545 DATNGSDLFHFAADSGIIPVYTLANGKQYVAVEVG 581

RESULT 6
DHET_ACCEU STANDARD; PRT; 742 AA.
AC P18278;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RA MEDLINE=89255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RA Yano K.;
RA "Cloning and sequencing of the gene encoding the 72-kilodalton
RA dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RA aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RP [2]
RP 3D-STRUCTURE MODELING.
RA MEDLINE=9528964; PubMed=7772016;
RA Cozler G.E., Giles I.G., Anthony C.;

```


FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT TURN 146 148
 FT HELIX 149 149
 FT STRAND 151 151
 FT STRAND 157 159
 FT TURN 160 161
 FT STRAND 162 166
 FT STRAND 168 168
 FT HELIX 170 172
 FT TURN 173 173
 FT STRAND 177 182
 FT TURN 183 185
 FT STRAND 188 193
 FT TURN 198 202
 FT STRAND 203 203
 FT TURN 205 210
 FT TURN 212 213
 FT HELIX 219 222
 FT TURN 223 223
 FT HELIX 226 231
 FT STRAND 241 243
 FT TURN 244 247
 FT STRAND 248 252
 FT HELIX 261 263
 FT TURN 269 270
 FT STRAND 273 278
 FT TURN 279 281
 FT STRAND 284 289
 FT STRAND 305 311
 FT TURN 312 313
 FT STRAND 314 322
 FT TURN 324 325
 FT STRAND 327 332
 FT TURN 333 335
 FT STRAND 336 343
 FT TURN 346 347
 FT STRAND 351 355
 FT TURN 356 359
 FT STRAND 360 363
 FT HELIX 365 367
 FT TURN 370 370
 FT STRAND 371 372
 FT STRAND 373 373
 FT STRAND 375 378
 FT TURN 382 383
 FT STRAND 390 390
 FT STRAND 392 394
 FT TURN 395 398
 FT STRAND 399 405
 FT STRAND 407 413
 FT TURN 420 421
 FT STRAND 427 433
 FT TURN 435 436
 FT TURN 438 439
 FT STRAND 444 450
 FT TURN 451 454
 FT STRAND 455 462
 FT STRAND 469 469
 FT STRAND 471 473
 FT TURN 474 476
 FT STRAND 477 481
 FT TURN 483 484
 FT STRAND 486 491
 FT TURN 492 495
 FT STRAND 496 502
 FT STRAND 512 516
 FT TURN 517 518
 FT STRAND 519 526

FT TURN 530 533
 FT HELIX 534 538
 FT TURN 539 539
 FT TURN 543 545
 FT HELIX 546 552
 FT TURN 553 555
 FT HELIX 556 558
 FT TURN 559 559
 FT STRAND 565 570
 SQ SEQUENCE 571 AA; 62449 MM; 0BE94EA5AD2AB1E1 CRC64;

Query Match 16.4%; Score 332.5; DB 1; Length 571;
 Best local similarity 25.1%; Pred. No. 2,3e-18;
 Matches 111; Conservative 65; Mismatches 161; Indels 105; Gaps 18;

QY 3 ANGVIVA-----GSTGQYSPF-----GC-----FVSGHD 26
 DB 123 ANGHLLALDAKTGINWEVEVCDPKYGSFTLQAPFAKDTVLMSGAGELVGAVNAFD 182
 QY 27 SATGEELMR-----NYFIPRAGE--EGDETGNQDYEARMTGA---WQ 65
 DB 183 LKTGELKWRAPATGSDSVRLANDFNSANDHYGFGIGKTWEGD--AMKIGGTMGW 239
 QY 66 ITYDPYTNLVHYGSTAVGASETQGTGPG-----TLVGNTRFAVRPPTGELVHRHOTL 120
 DB 240 YAYDPKLNLFYQSSGNPAPNENETMR--PGDNKKTMTIMGRDL-----DTGAKKQYQKT 291
 QY 121 PRDNMOECTFEMMTATNDVOP--STEMEGLQS--INPNAATGERRVLTGVPCKTGIMQPD 178
 DB 292 PHDEMDFAGVNGVLTLD--QPVNAKTFPLSHIDRN-----GLYTLN 332
 QY 179 AETGEFLMARDTN--YQNMIESIDENGIYVNEALIKELDVE--YDVCPTFLGRDMPSPA 236
 DB 333 RENGNIIVAEKVDPAVAVRKKVDLKTGTPVRODEPATRMDHGTNCPSPAMGHNGVDS 392
 QY 237 LNPDSGIVFPLNNWCYDMAVAODEFTSMDEVYTSNVTLP---PGKMDIGRIDADIS 292
 DB 393 YDPESRTLVAGLNHICMDNEPFLPYRAGQFPGALITAMPNGNGTKRMEQIRAFDIT 452
 QY 293 TGRITMSVERAANYSPIVSTGGGVLFNGGTDRIYFALSOEGETIMOTRLATVASQAI 352
 DB 453 TGRAKWTWKEKPAWGGTLTYTGGLVWATLDDGYLKALDNKDKELMKNKMSGGIGSPM 512
 QY 353 SYEVDGMOYVAIAGGVSYSG 374
 DB 513 TYSFKGQYI-----GSMGVG 529

RESULT 8
 DHET_ACEPO STANDARD; PRT; 738 AA.
 AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -I- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +

```

CC      -1- COFACTOR: POQ AND HEME.
CC      -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC      SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC      COMPLEX.
CC      -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC      SPACE (POTENTIAL).
CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed, usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL; D00635; BA00528.1; -.
CC      PIR; S14270; S14270.
CC      HSSP; Q9Z4J7; 1FLG.
CC      InterPro; IPR001479; Bac_PQO.
CC      InterPro; IPR002372; Bac_PQO_repeat.
CC      InterPro; IPR000345; CytC_heme_bind.
CC      Pfam; PF01011; Bacterial_POQ_6.
CC      PROSITE; PS00363; BACTERIAL_POQ_1.
CC      PROSITE; PS00364; BACTERIAL_POQ_2; 1.
CC      PROSITE; PS00190; CYTOCHROME_C; 1.
CC      Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
CC      KW SIGNAL 1 35
CC      FT CHAIN 36 738
CC      FT BINDING 650 650
CC      FT BINDING 653 653
CC      FT METAL 654 654
CC      FT SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92A66 CRC64;
-----
Query Match 16.4%; Score 332.5; DB 1; Length 738;
Best Local Similarity 22.8%; Pred. No. 3,2e-18;
Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;
-----
QY      2 VANGVIAGSTQCYSPFGC-FVSGHSDATGELMKRNYFIPRGERGCD----- 47
DB      204 IAKGRVITGN-GGSDFGARGFVSADFADETGRKVDREFFVPPMKNEPDDAASDYLANKAY 261
QY      48 EHWGMDYERKMTG-----WGQITTPYNTLVHVGSTAVGPASETGRTPGCTL 97
DB      262 QTWS-----PTGAMTRGGGGGVWDSTIVYDVADLVYLGNGSPWNYKYTRSECKGN 314
QY      98 YGTNTRFAVRPTGTGIWHRQTLPRDPMWDECTFEFMAVTVNVQVSTENEGLOSINPNA 157
DB      315 LFLGSLVALKPKETGEVWHFOETPRDQMFTSDQIMTLDEL----- 357
QY      158 TGERRYLTGVPCKTGTMMQFDAETGEFFLARDNTIONMTESID-ENGIVTNEDALKEFL 216
DB      358 NQETRHVIYHAKRNGFFYLIDAKTGEFFISGKNVYVYVWMASSGDPKGRPIYNDALYTLT 417
QY      217 DVEYDCPTFLGLGRMPSPALPDGSIYRPLNNVCYDMMAAVDQETT-EMDYVNTS---N 272
DB      418 GKEMVGIPEDDLGHNFAAFAAFSPKTYLYIIPAQVFFLTTNOVGSTPTPRHDSNNLGLDAN 477
QY      273 VTKLPG-----KDMIGRIDAIDISTGRITMSVERAANSPVLSITGGGVLFNGGND 324
DB      478 KVGIPDSPRAKQAFVYDKLGKVIYAMPDKQAEAMRVYDHGPNNGGILATGGLFQGLAN 537
QY      325 KTFRALSETGETIMOTRLATYASGAISIVYDGMQVYVLAG 367
DB      538 GEHFAYDATNGSDLFHFAADSGIAPVYLLANGKQIVAVEVG 560
-----
RESULT 9
EXAA_PSEAE STANDARD; PRT; 623 AA.
ID EXAA_PSEAE
AC Q9Z4J7;

```

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (DEHD).
 GN EXAA OR PA1962.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17933;
 RA Diehl A., Wiltzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme."
 RL Eur. J. Biochem. 257:409-419(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RA Stover C.K., Pham X.-O.T., Errin A.L., Mizoguchi S.D., Warren P.,
 RX MEDLINE=20437337; PubMed=10984043;
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gatter R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 RL [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=99173751; PubMed=10075429;
 RA Schoebert M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase."
 RL Microbiology 145:471-481(1999).
 RL [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Ketzel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT Pseudomonas aeruginosa: basis of substrate specificity."
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: PO4 AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PO4 DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: A1009858; CA008896.1; -;
 DR EMBL: AE004624; AAG05370.1; -;
 DR EMBL: AF068264; AAC79657.1; -;
 DR PDB: 1FLG; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PoQ.
 DR InterPro: IPR002372; Bac_PoQ_repeat.
 DR Pfam: PR01011; Bacterial_PoQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.

DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Periplasmic; Signal; Calcium; 3D-structure;
 KM Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623 QUNOINPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;

Query Match 15.2%; Score 308.5; DB 1; Length 623;
 Best Local Similarity 22.9%; Pred. No. 1.9e-16;
 Matches 88; Conservative 68; Mismatches 175; Indels 53; Gaps 10;

QY 26 DSATGEELMNNYFPR-----AGEEDETWGNDEA-----RMTGA--WG 64
 DB 224 DPTGEITMRRPVEGGMRLNGSKDSTVTGDKVAKSPMDRNSPTGKVESWSHGGAPMQ 283
 QY 65 QITVDPYTNLVHGSTVAVPASETQRTGPG-----TLTGTWTRAVRPDGEIYWRHQ 118
 DB 284 SASFDAETNTIIVAGAGPMPWNTWARTAKGNPHDSDLT-ITSGQVDPDSSEVAKWFTQ 342
 QY 119 TLRDWMDOCTEEMAVTVNDVOPSTEMEGLSINPN-----AATGERRVLGVPCKGT 173
 DB 343 HTFNDAWDESGNNEVLFDYKAKDKIKVATAHADNNGEYVYDRSNGKLQNAFPVDNI 402
 QY 174 MW--OFDAETGEFLMARDTNYQNMIESIDENGIVTVNEDAILKELDVEYDVCPTFLIGRD 231
 DB 403 TMSHIDDKTG-----RPEREGQRPPLEPG-----QKHGKAVEVSPPLFGGKN 447
 QY 232 WPSAALNPDSGIYFIPLNNVCYDMAVDOEFTSMVYNTSNYTKLPPGKMDIRDAIDI 291
 DB 448 WNMMAVSODTGLFVYVPAHMKEDYWEVYSYTKGSAVILMGFRIKMYDHDHGLSRAMP 507
 QY 292 STGRTLSVERAANSPVLTSGGVLNFGNDRIYFRALSOETGELTQTRATYVASGQA 351
 DB 508 VSKVYMEHKEHLPLWAGVATAGNLVFTGTDGDFKAPDASGKELMFKTGSGIVSPP 567
 QY 352 ISEYVDMQYVAIAGGVSYSGL 375
 DB 568 ITWEDGBOYLGYT---VGTGGAV 588

RESULT 10
 DHET_GLUOX STANDARD; PRT; 757 AA.
 AC 005542;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).
 GN ADHA.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconobacter.
 NCBI_TaxID=442;
 RN 11
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
 RC STRAIN-IFO 12528;
 RX MEDLINE-97208225; PubMed-9055427;
 RA Kondo K., Horinouchi S.;
 RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in *Acetobacter pasteurianus*."
 RL Appl. Environ. Microbiol. 63:1131-1138(1997).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC

CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D86375; BA019753.1; -.
 DR HSSP: Q9Z4U7; IFCG.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT MOD_RES 35 35 PYRROLIDONE CARBOXYLIC ACID.
 FT ACT_SITE 342 342 BASE (POTENTIAL).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82968 MW; 39B9F903B947581 CRC64;

Query Match 14.7%; Score 298.5; DB 1; Length 757;
 Best Local Similarity 23.2%; Pred. No. 1.5e-15;
 Matches 99; Conservative 68; Mismatches 180; Indels 79; Gaps 13;

QY 2 VANGIYVAGSTQYSPGCG--FVSGHDSATGEELMNNYFIPRAGEED----- 47
 DB 202 IAKGKVLGN--GGAERFAGSFSAFDERSKLDWRFYVPPNPKKDGAAASDIILSKA 259
 QY 48 -ETWGNDEYFARMGTG---AMGQITVDPYTNLVHGSTVAVPASETQRTGPGTLTGNTNR 103
 DB 260 YPTGKNGAMKQGGGGGVMDSLVYDPTDLVYLGNGSPMNKFFSEBKGNNLPLIGST 319
 QY 104 FAVRPDGEIYWRHQITLPRDWMDOCTEEMAVTVNDVOPSTEMEGLSINPNATGE-RR 162
 DB 320 VAINPDGKYVWHEQETPMDEMDTYSVQIINTLDMVY-----NEMER 362
 QY 163 VLTGVPCKTGTMMQFDAETGEFLMARDTNYQNMIESIDE-NGIVTVNEDAILKELDVEYD 221
 DB 363 VIVHAP-KNGEFYIIDAKTKGFTGKRYTEENANGLDPTGKRPYVPDLMTLTGKPMV 421
 QY 222 VCPFLGGRDMPSAALNPDSGIYFIPLNNV-----CYDMAVDOEFTSMVY 268
 DB 422 GIPBELGLGHNFAMAYSPKTKLVIPAOQIPLLYDQGGKFAYHDAMNGLDMNKIGLF 481
 QY 269 NTSNVTKLPKGMIGRIDALIDISTGRTL-WSEVERAANNS-----PVLTSGGVV 318
 DB 482 DDNDDEHVAANKDF-----LKVLGKVTWAMDEKNAAPATTINHKPKPMNGGLATAGNYI 535
 QY 319 FNGSTDRYFRALSOETGELTQTRATYVASGQAISEYVDMQYVAI-----AGGG 368
 DB 536 FQGLANEFILAYATNGNDLISFPASAIITAPVYTTANGKQYVAVEVGMGIFPLVGG 595
 QY 369 VYSGSG 374
 DB 596 VARTSG 601

RESULT 11
 DHGA_ACICA STANDARD; PRT; 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)


```

01-NOV-1988 (Rel. 09, Last sequence update)
DF 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDH-A.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxId=471;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=LMd 79.41;
RX MEDLINE=88289368; PubMed=3399393;
RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
RT "Nucleotide sequence of the gene coding for quinoprotein glucose
dehydrogenase from Acinetobacter calcoaceticus."
RL Nucleic Acids Res. 16:6228-6228(1988).
CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERS
CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
CC DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
CC EMBL, X07235; CAA30222.1; -.
CC PIR, S00943; S00943.
CC DR InterPro: IPR001479; Bac_POQ.
CC DR InterPro: IPR002372; Bac_POQ_repeat.
CC DR Pfam: PF01011; Bacterial_POQ_7.
CC DR PROSITE: PS00363; BACTERIAL_POQ_1.
CC DR PROSITE: PS00364; BACTERIAL_POQ_2.
CC KM Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
CC FT SIGNAL 1 33 POTENTIAL.
CC FT CHAIN 34 801 GLUCOSE DEHYDROGENASE-A
CC FT [PYRROLOQUINOLINE-QUINONE].
CC FT TRANSMEM 39 55 POTENTIAL.
CC FT TRANSMEM 59 79 POTENTIAL.
CC FT TRANSMEM 94 108 POTENTIAL.
CC FT TRANSMEM 119 138 POTENTIAL.
CC FT ACT_SITE 471 471 BASE (POTENTIAL).
CC SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;

```

```

Db 506 LDRR-NGQPIVPTKRPVPTQVTKRGQOTKEFYSKTPQPSDNLMLAQDRLKTRKDMWGAT 5633
QY 188 -----RDTRNYOMNISIDENGLVYVNEADAILKELDEYDQVPTFLGGSDPMSAA 2366
Db 564 MDJOLMCVRSEKRLNBYDITYPPSENGTL-----YPPGNLGFTEWGMGS 607
QY 237 LNPOS-----GIYFI-----PLNNYCYDMAVDOEFTS 264
Db 608 VNPDRQAVAMNPICGLPEFSRLIPADPNNAQTKAGAGTEGQGVQPYGVBY-----G 657
QY 265 MDVYNTSNVTKLPPEKDMIGRIDAILDISTGRTLM-----SYERAAN 306
Db 658 VEISAFSLPLGLPCKQPMAGVYAGVDILKTHVYVWKKRIGTIRDSLPLNLFOLPAVKIGVPG 717
QY 307 YSPVLSTGGGVLNFNGCT--DRFRALTSQETGETLWOTRLATVAASGA--ISYEVGMQYVA 363
Db 718 LGGSTLSTGNWMEFATODNYLRAFNVTNGKKLMEARLP--AGQQAIPMTYEINGKQYV 775
QY 364 I-AGGVSYS 373
Db 776 IMAGHGSGFGT 786

RESULT 12
DHG_ECOLI STANDARD; PRT; 796 AA.
ID P15877;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase [pyriroloquinoline quinone] (EC 1.1.99.17).
GN GCD OR B0124.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91035240; PubMed=2228962;
RT Cleon-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
RT "Cloning, mapping, and sequencing of the gene encoding Escherichia
RT coli quinoprotein glucose dehydrogenase.";
RL J. Bacteriol. 172:6308-6315(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RX MEDLINE=93123180; PubMed=8419307;
RA Yamada M., Asaoka S., Sailer M.H. Jr., Yamada Y.;
RT "Characterization of the gcd gene from Escherichia coli K-12 W3110
RT and regulation of its expression.";
RL J. Bacteriol. 175:568-571(1993).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RX MEDLINE=94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 2.4-4.1 min (110,917-193,663 bp) region.";
RL Nucleic Acids Res. 22:1637-1639(1994).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [5]
RN TOPOLOGY.
RX MEDLINE=93286127; PubMed=8509415;

```

RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.,
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT Escherichia coli and its ubiquinone-binding site.",
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Cozler G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of Escherichia
 RT coli modelled on that of methanol dehydrogenase from Methylobacterium
 RT extorquens";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51323; CA35706.1; -
 DR EMBL: D12651; BA02174.1; -
 DR EMBL: D26562; CAB20298.1; -
 DR EMBL: AE000122; AAC73235.1; -
 DR PIR: JY0107; JY0107.
 DR HSSP: P38539; 4AAH.
 DR Ecogene: EG10369; gcd.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR Oxioreductase; POQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 37 37 PROBABLY.
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLY.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLY.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 96 110 PROBABLY.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLY.
 FT DOMAIN 142 796 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TUSADATP -> HIKRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E5 CMC64;

Query Match 11.4%; Score 231.5; DB 1; Length 796;
 Best Local Similarity 23.4%; Pred. No. 2.5e-10;
 Matches 113; Conservative 53; Mismatches 154; Indels 163; Gaps 23;

QY 6 VIVAGS-TCOXS--PFGFVSGHDSATSEELMRNFIRPRA-----GEEDDFWGDNDYE 55
 DB 346 IYMASVVDNSTRSTRETSCVIRKFDVNTGELLMA--FDGADDPNAPISDEHTFFNS--- 400
 QY 56 ARMTGANGQITDPVFNILVHYGSTAVGPASETORGTPGGLTGTNTRFAVRPTGELIYW 115
 DB 401 ----PNSWAPAPAYDAKDLVLLPVGVTTPDIMGGRNRPDEGRVASSI-LALNATTKGLAW 455

QY 116 RHQITLPRDMDQCEPFEAMVTVNDVQPSSTEMGLSINPNNAATGRRVLTGVPCKTGTMW 175
 DB 456 STQTVHHDMDMD-----LPAQPT-----LADITVN-----GQKVVIYAPAKTGNI 498
 QY 176 OFDAETGEFL-----WARDNYOMMIESIDENGI-----VYVNEDAI 212
 DB 499 VLDNRNGELVYAPKPEPVYQGAARDYTPPOPESELSFRPTKLSGADMGATFQDLV 558
 QY 213 LKEL-DVEYD-----VCPTPLGRDWSAALNP----- 239
 DB 559 CVMFHMRYEGIFTPPSBOGLTFPPNLMGMEWGISVDDPREVALIANPALPFSKLI 618
 QY 240 -----DSGI-----YFIPLNVCYDMAVDOFTSMDYVNSNV 273
 DB 619 PRGPGNPMPQPDANGKTGESGIPQYGVPPVGTLTN-----PFLS-----P 659
 QY 274 TLPPGKDMIGRIDAIDISTGRTLWSVERAANYS-----PV-----LSTG 314
 DB 660 FELPKOPAMGYISALDLKTNEVYWKKRIGTPQDSMPPEMPPVPFNNGMPLGPISTA 719
 QY 315 GGVLENGGT-DRIYFALSQETGELIMQTRATVASGGA--ISYEVDQMUYAI-AGGGSV 370
 DB 720 GNVLEIATADNYLRAVNMNSGEKIMQGRLP--AGGATPMTYEVNGKQYVIVASGHS 777
 QY 371 YGS 373
 DB 778 FGT 780

RESULT 13
 QYIA-ACICA
 ID QYIA-ACICA STANDARD; PRT; 809 AA.
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [Pyroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN QYIA.
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1.
 RX MEDLINE-9505936; PubMed-8002591;
 RA Elsemore D.A., Ornstom L.N.;
 RT "The pca-pob supraperoic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE-96011389; PubMed-7592351;
 RA Elsemore D.A., Ornstom L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: Quinate + pyroloquinoline-quinone = 5-
 CC dehydroquinate + reduced pyroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: POQ.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PYROLOQUINOLINE-QUINONE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- INDUCTION: BY PYROLOQUINOLINE-QUINONE.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L05770; AAC37161.1; -

CC InterPro: IPR001479; Bac_POQ.

CC InterPro: IPR002372; Bac_POQ_repeat.

CC Pfam: PF01011; Bacterial_POQ_7.

CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.

CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.

CC Oxidoreductase; POQ; Quinate metabolism; Transmembrane.

CC

CC TRANSMEM 14 34 POTENTIAL.

CC TRANSMEM 41 61 POTENTIAL.

CC TRANSMEM 68 88 POTENTIAL.

CC TRANSMEM 90 110 POTENTIAL.

CC TRANSMEM 127 147 POTENTIAL.

CC SEQUENCE 809 AA; 88196 MW; 71P67CEBA62BFCB CRC64;

Query Match 10.1%; Score 204.5; DB 1; Length 809;

Best Local Similarity 22.9%; Pred. No. 3,1e-08;

Matches 105; Conservative 54; Mismatches 187; Indels 113; Gaps 17;

QY 2 VANGVIVAGSTCYSPF-----GCFVSGHDSATGELMKNYFIPRAGEGDEFTWGDYAR 57

DB 354 IAGTIVSGRIADNVADMPGIVRAYDVTIGKLRMA--EDPRNDPNVYKLPGEIKR 411

QY 58 WMTGAMGOITVDPVNLVH--VGTAVGPASETORGCTGTYGNTFFAVAPDGEIWM 115

DB 412 SSTNMAASSTDPOMATVFLPMGSSVDYWGGR--TAADHYNTSV--LALDATGKREK 468

QY 116 RHQTLEPRDMDOECTFEMVATVNDVQSTEMEGLQSIENFNAATGERVLTGPKCTGTW 175

DB 469 VNTVAVNDLMD-----FDLPMPQS-----LVDFPKDGTTPRAVYIGT--KSGOFY 512

QY 176 QPDAETGEFL-----WADPTNYOMIESIDENIVYNE-----DAI 212

DB 513 VLDRTVKGKPLRYIEPIKADIPGEYKSTQPRVEMQIQNQLKESDMGAPFPOL 572

QY 213 LKELD---VEYD-----VCPFLGGRDPSALNPDGTYF----- 245

DB 573 MCRINKSKRYDGLYAPGTDSLSFPGSLGGMNGSIAFDTHRYMFVNDRLGLMTQL 632

QY 246 -----IPLN--NVCYDMNAVDEFTSMQVNTSVTKL--PPGKDMIGRIDAIDI 291

DB 633 IKQTPEDIKIOANGGEKVTMGMAVPMKGTPIKVNKNRPMASLGIIPCOKPPFGTMTAIDM 692

QY 292 STGRITLMSVERAANS-----PVLSTGGGVLFGNGDRYFRALS 331

DB 693 KTRQVAVQVPLGTIDTDPGIMKIGLAKIDGIMPTIGGPATOGGLVFAATDYLIRAFN 752

QY 332 QETGETLMOTRLATVASGOAISY--EVDGMQVYVAINGGG 368

DB 753 SSGKELMKARLPVSGGTPTMSPKTKQYVVSAGG 791

RESULT 14

DHG_GLUOX STANDARD; PRT; 808 AA.

AC P27175;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor

DE (EC 1.1.99.17).

GN GH.

OS Gluconobacter oxydans (Gluconobacter suboxydans).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

OC Gluconobacter.

OX NCBI_TaxID=442;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-92017653; PubMed-1833618;

RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;

RT "A single amino acid substitution changes the substrate specificity

RT of quinoxaline glucose dehydrogenase in Gluconobacter oxydans";

RL Mol. Genet. 229:206-212(1991).

RN [12]

RP REVISION TO 213.

RA Goosen N.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +

CC reduced acceptor.

CC -1- COFACTOR: POQ.

CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.

CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE

CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN

CC HERE.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X62710; CAN44594.1; ALT_SEQ.

CC PIR: S17716; QPKEX.

CC InterPro: IPR001479; Bac_POQ.

CC InterPro: IPR002372; Bac_POQ_repeat.

CC Pfam: PF01011; Bacterial_POQ_7.

CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.

CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.

CC Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.

CC

CC KW SIGNAL 1 33 POTENTIAL.

CC FT CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-

CC QUINONE].

CC FT TRANSMEM 35 54 POTENTIAL.

CC FT TRANSMEM 59 76 POTENTIAL.

CC FT TRANSMEM 94 108 POTENTIAL.

CC FT TRANSMEM 123 138 POTENTIAL.

CC FT ACT_SITE 470 470 BASE (POTENTIAL).

CC FT VARIANT 788 788 H -> N (IN P2 FORM).

CC SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match 9.9%; Score 200.5; DB 1; Length 808;

Best Local Similarity 20.9%; Pred. No. 6,4e-08;

Matches 101; Conservative 61; Mismatches 173; Indels 149; Gaps 19;

QY 1 IVANGVIV-AGSTOYSPFCFVSGHDSATGELMKNYFIPRAGEGDEFTWGDYAR 55

DB 349 ITANALINDNSVKAQAS--ATOAFDYITKRWVWDASPPNDQLPDESHVYFHNNSP 405

QY 56 ARWMTGAMGOITVDPVNLVH--VGTAVGPASETORGTP--GGTLYSTNTRF-----AVR 107

DB 406 NSWIVS-----SIDANLNLVYIPMGV-----GFPDQGGSDRTKDSERFAGIYALN 451

QY 108 PDTEIVMRHQTLPDMNDQECTFEMVATVNDVQSTEMEGLQSIENFNAATGERVLTG 167

DB 452 ADTGKLAWFYQTVHDDMDLPSQSLVDYTKDGLVPAIYV----- 495

QY 168 PCKTGTMWQPAENG-EFLMARDT-----NYONNIESIDENGLATVVEDAILKELD 217

DB 496 PKTDIDIVLDRKKEIVPAPETVPVPGGAAPGHTSPTQMSQ--LTLAKRPNLNDSD 552

QY 218 V-----EYD-----VCPFLGGRDPSALNPDGTYFIP 247

Db 553 IWGCTIDDMFCSTFYHTLNEGFPFPPSLKSLIFPPGLDMFEGWGLAVDPROYAFAN 612
 QY 248 LNNVCYDMAVADQ-----EFTSMQVYNTSNTKLPKDKMI----- 283
 Db 613 PISLPEFSQSLVPRGPGNLMPEENAKGTGSETGLQHNHYGIPYAVNLHPLFDPLLPGLK 672
 QY 284 -----GRDAIDISTGRTLSVERAANYS-----PV-----LSTG 314
 Db 673 MPCRTPMGVYAGIDLTNNKVVQHRNGTLRDSMYSLLPPLPIKIGVPSLGGPLSTA 732
 QY 315 GGVLF-NGTDRYFRALSOETGELMOTRLATVASGAISYEVDGMQYVALIAGGVSYS 373
 Db 733 GNLGFLTASMDYITRAVNLTTGKVLMDRLPAGAQAATPIYAINKQYI-----VTYAG 786
 QY 374 GLNS 377
 Db 787 GHNS 790

RESULT 15
 ID QYIA_XANCU STANDARD; PRT; 790 AA.
 AC Q9XD78;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable quinolate dehydrogenase [Pyrroloquinoline-quinone]
 DE (EC 1.1.99.25).
 GN QUMA.
 OS Xanthomonas campestris (pv. Juglandis).
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xanthomonas.
 ON NCBI_TaxID=4291;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C5;
 RX MEDLINE=20063481; PubMed=10594704;
 RA Lee Y.-A., Lo Y.-C., Yu P.-P.;
 RT "A gene involved in quinolate metabolism is specific to one DNA homology
 RT group of Xanthomonas campestris.";
 RL J. Appl. Microbiol. 87:649-658(1999).
 CC -1- CATALYTIC ACTIVITY: Quinolate + pyrroloquinoline-quinone = 5-
 CC dehydroquinolate + reduced pyrroloquinoline-quinone.
 CC -1- COFACTOR: PQQ (BY SIMILARITY).
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF109471; AAC38453.1; -
 DR InterPro; IPR001479; Bac_PQQ.
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ; 6.
 DR PROSITE; PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE; PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 KW Oxidoreductase; PQQ; Quinolate metabolism; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 77 94 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 SQ SEQUENCE 790 AA; 82896 MW; 875F29B52A49FE6F CRC64;

Query Match 8.9%; Score 180.5; DB 1; Length 790;
 Best Local Similarity 22.3%; Pred. No. 2.2e-06;
 Matches 93; Conservative 43; Mismatches 150; Indels 131; Gaps 17;

QY 51 GNDYERAMWTGAMGQITTYDPTNLVHYSSTAVGASSETQRTPGGTLTYGN----- 101
 Db 386 GSSY-VRSTPNWAMPSYDAAMNTVF-----LPLGGPSTDIYGAERTALDHRYG 433
 QY 102 -RFAVRPDGTGEIVRHOITLPRDNDQCTEEMANTNDVPSPEMESLQSIINNAATGE 160
 Db 434 ASVALDATTGAEKVVYQTVHNDLMD-----FDLPQPSL-----IDEPNODGS 477
 QY 161 RRVLTGVPCKTGTMMQFPAETGEFLW-ARDT-----NY-----QNMIESIDENGIVT 206
 Db 478 HTPAVVIGTKAGQIYVLBRATGKPLTEVREVPVAGSDIAHQVAPTPQPLSGMQITKH 537
 QY 207 VNEDAILKEIDVEYDVC-----PTFLGGRDPSAALNDPSG 242
 Db 538 LTESDMGATAMDQMLCRIAFKOMREGLYTPGTDVSLSPGSLGGMNMGSLSTDVHD 597
 QY 243 IYF-----IPLNN-----VCYDMAVDEFTSMQVYNTSNTKLP--- 276
 Db 598 VVFANDMRLGLMVQMTIPADTRKAPDAAGGEAVNTGKGVPLKGPYAVNNRFLSLALGIP 657
 QY 277 ---PPGKDMIGRIDAIDISTGRTLSVERAAN-----YSPV-----LSTG 314
 Db 658 COAPP---YGTLSAIDIKTRSIAMQVPVGVQDTGPFPGIKMHLPIPIGMPITLGGTISTQ 713
 QY 315 GGVLFNGT-DRTYFRALSOETGELMOTRLATVASGAISY--EVDGMQYVALIAGG 368
 Db 714 GGLVFIACTQDYLRAFDASATGKELMKGRPLVVGSGGPITYVSHKTGQYVVISAG 770

Search completed: May 24, 2002, 10:15:52
 Job time: 770 sec

Fri May 24 11:27:11 2002

us-08-934-506a-5_copy_180_556.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:32 ; Search time 93.93 Seconds

(without alignments)
694.337 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556

Sequence: 1 IVANGYIVAGSTCOYSPFCG.....GMQYVAIAGGVSYSGSLNS 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mmc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertibrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriaph:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1275.5	62.8	608	Q93RE9	Q93RE9 pseudogluc
2	394	19.4	698	Q9KH03	Q9KH03 alcaligenes
3	380.5	18.7	601	P71509	P71509 methylobact
4	378.5	18.6	629	Q9A048	Q9A048 methylobact
5	375	18.5	601	Q9EYW8	Q9EYW8 rhizobium m
6	374	18.4	601	Q92WY9	Q92WY9 rhizobium m
7	360.5	17.7	695	Q9F9U2	Q9F9U2 pseudomonas
8	359	17.7	599	Q9L935	Q9L935 pseudomonas
9	355.5	17.5	708	Q46444	Q46444 comamonas t
10	350.5	17.2	633	Q24759	Q24759 hyphomicrob
11	338	16.6	742	Q53362	Q53362 acetobacter
12	335.5	16.5	573	Q59540	Q59540 methylomphal
13	323	15.9	695	Q934G0	Q934G0 pseudomonas
14	299.5	14.7	691	Q9AF95	Q9AF95 pseudomonas
15	281.5	13.9	623	Q9AGW3	Q9AGW3 pseudomonas
16	226.5	11.1	790	Q9X2S5	Q9X2S5 pantoa cit

17	208.5	10.3	182	2	Q32699	Q32699 hyphomicrob
18	204.5	10.1	470	2	Q30326	Q30326 acetobacter
19	203.5	10.0	179	2	Q32697	Q32697 hyphomicrob
20	200.5	9.9	180	2	Q32696	Q32696 hyphomicrob
21	200.5	9.9	180	2	Q32700	Q32700 hyphomicrob
22	200.5	9.9	181	2	Q32615	Q32615 hyphomicrob
23	199.5	9.8	181	2	Q32692	Q32692 hyphomicrob
24	197.5	9.7	181	2	Q32703	Q32703 hyphomicrob
25	196.5	9.7	181	2	Q32621	Q32621 hyphomicrob
26	195.5	9.6	182	2	Q32706	Q32706 hyphomicrob
27	193.5	9.5	180	2	Q32701	Q32701 hyphomicrob
28	191.5	9.4	182	2	Q32612	Q32612 hyphomicrob
29	191.5	9.4	184	2	Q32693	Q32693 hyphomicrob
30	190.5	9.4	172	2	Q33894	Q33894 hyphomicrob
31	186.5	9.2	172	2	Q33882	Q33882 methylotinu
32	186.5	9.2	179	2	Q32707	Q32707 hyphomicrob
33	185.5	9.1	180	2	Q32694	Q32694 hyphomicrob
34	185.5	9.1	180	2	Q32704	Q32704 hyphomicrob
35	185	9.1	171	2	Q33884	Q33884 methylotinu
36	183.5	9.0	179	2	Q32702	Q32702 hyphomicrob
37	182.5	9.0	184	2	Q32695	Q32695 hyphomicrob
38	181.5	8.9	185	2	Q918K8	Q918K8 uncultured
39	177	8.7	644	2	Q52551	Q52551 pseudomonas
40	176.5	8.7	172	2	Q93K58	Q93K58 methanotrop
41	175.5	8.6	172	2	Q93J25	Q93J25 methylotell
42	174.5	8.6	179	2	Q32698	Q32698 hyphomicrob
43	174	8.6	803	16	Q91115	Q91115 pseudomonas
44	172.5	8.5	185	2	Q918L0	Q918L0 uncultured
45	169.5	8.3	185	2	Q918K9	Q918K9 uncultured

ALIGNMENTS

RESULT 1
ID Q93RE9 PRELIMINARY; PRT; 608 AA.
AC Q93RE9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharoketogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shibata T., Saito Y.;
RT "Alcohol dehydrogenase";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1; -
SQ SEQUENCE 608 AA; 65101 MW; 0ACBC97AE1BA570 CRC64;

Query Match	62.8%	Score 1275.5;	DB 2;	Length 608;
Best Local Similarity	60.4%	Pred. No. 2.6e-88;		
Matches 236;	Conservative 43;	Mismatches 97;	Indels 15;	Gaps 3;
QY 1	IVANGYIVAGSTCOYSPFCGVSCHDSATGELMRNFIPRAGEGDETMGN-DYEARM 59			
DB 207	IVANGYIVAGSTCOYSPFCGVSCHDSATGELMRNFIPRAGEGDETMGN-DYEARM 266			
QY 60	TGAGQITTYDVTNLMVHGSTAVGPASETORGTGTLGNTNFAVAPDPTGLIWRHQT 119			
DB 267	TGAGQITTYDVTNLMVHGSTAVGPASETORGTGTLGNTNFAVAPDPTGLIWRHQT 326			
QY 120	LPRNMDOCTFFEMAVTNVDYOSTEMEGLOSTIPNATGE-RVLTGVCCKTGTMQFD 178			
DB 327	LPRNMDOCTFFEMAVTNVDYOSTEMEGLOSTIPNATGE-RVLTGVCCKTGTMQFD 386			
QY 179	AETGEPLMARDTNYQNMIESIDENGIVTVNEDATLKLDEYDVDCPTFLGRDMPASALN 238			

```

DB 387 AKTGDFWMSKATVEONSISIDDTGLVTVEMDLKEPKETNYCPTFLGGRDMPAGYL 446
QY 239 PDGIEFIPLNANCYDMANADDEFMSDYNNTSNKLPKGMIRIDALIDISTERTM 298
DB 447 PKSNLYIPLSNACYDMATTEATPADYNTDAILVLPKGMIRIDALIDISTERTM 506
QY 299 SVERAANYSPLVSTGGVLFNGGIDRYFRALSOETGELTMOTRLATVASGOAISYEVDG 358
DB 507 SYEFRAALYDPLVLTGGDLVFGVIGIDRFRALDAESGKSWTRLPAGAVSGYTSIDG 566
QY 359 MOVVALANGG-----VSYGSGIN 376
DB 567 ROYVAVSGSLGPGTPTPDVDSAGAN 597

RESULT 2
Q9KH03 PRELIMINARY; PRT; 698 AA.
AC 09KH03;
DT 01-OCT-2000 (Tremblrel, 15, Created)
DT 01-OCT-2000 (Tremblrel, 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE TERAHYDROFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_Taxid=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125557; Pubmed=11222593;
RA Zarni G., Schrader T., Andreessen J.R.;
RT "Catalytic and Molecular Properties of the Quinolomoprotein
RT Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo."
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL: AF273733; AAF6335.1; -.
DR HSSP: Q92477; IFIG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR00345; Cylc_heme_bind.
DR Pfam: PF01011; Bacterial_PQQ; 6.
DR Pfam: PF00034; Cytochrome_C; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180BD12FB2 CRC64;

Query Match 19.4%; Score 394; DB 2; Length 698;
Best Local Similarity 28.7%; Pred. No. 1.6e-21;
Matches 117; Conservative 61; Mismatches 161; Indels 68; Gaps 17;

QY 2 VANG-VIYAGSTCOYSPEGCFVSGHDSATGELMIRNYFP-----RAGEEGDETWG 51
DB 192 VYNGKYLIGNGAEYGVRS-YITAYDAETGKQOMRWYIVPGDPAFENEMAKAAATW- 249
QY 52 NNYEAR-WMTG-----AMGQITDPTVNLVHGSTAVGPASETORGTGG-ILYGTNTPFA 105
DB 250 -DPSKRYWINGGGVYVWMTMAPDEPLNMLYIGTGNAQFWSRKLRSKGDMLYLAASV-VA 307
QY 106 VAPDGEIYWRHQTLPBQWMOECFEEMVYN--VDQPSSTEMEGLSINNAATGERRV 163
DB 308 LMPDGEIYVWYQETPGDMDWTTSTODITLADLKIDGP-----RKV 349
QY 164 LTVGPKGTGMWQDAETGEFLMARDTNVQNMIESIDENG--IYTVNEDALIKELDEYD 221
DB 350 IHAAP-KNGFFVIDRTNGKFIKSAKNFYDVWMAISYDKNGRPVETPROADTSGKRA----D 404
QY 222 VCPFLGGRDMPAALNPDGSIYFPLNNVCYDMAVDEFTSMDYNTSN----- 272
DB 405 VVPGFGAHNMHSMFHKLAGIAPQHVPL-TLADNKEW-----VHNQKDSPEAHNGVG 459
QY 273 -----VTKLPKGMIRIDALIDISTERTMSVERAANYSPLVSTGGVLFNGGIDRY 326

```

```

DB 460 WNLGLVNAEPPRSKPMGRLLAMPDLAQKAVHRHDHAGPNNGTLATAGNLVFGGTADGR 519
QY 327 FRALSOETGELTMOTRLATVASGOAISYEVDGMOYVAIA-GGVSYSY 372
DB 520 LVAYHAATGEKLMQAPGSGVVAAPVYLLDGRQYVAVAGMGVYG 566

RESULT 3
P71509 PRELIMINARY; PRT; 601 AA.
AC P71509;
DT 01-FEB-1997 (Tremblrel, 02, Created)
DT 01-FEB-1997 (Tremblrel, 02, Last sequence update)
DE 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
GN MXAF.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_Taxid=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RX MEDLINE=97312011; Pubmed=9166622;
RA Christoserdova L., Lidstrom M.E.;
RT "Molecular and mutational analysis of a DNA region separating two
RT methyloctrophy gene clusters in Methylobacterium extorquens AM1."
RL Microbiology 143:1729-1736(1997).
DR EMBL: U72662; AAB58890.1; -.
DR HSSP: P38535; AAAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ; 7.
SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CB8239 CRC64;

Query Match 18.7%; Score 380.5; DB 2; Length 601;
Best Local Similarity 27.1%; Pred. No. 1.4e-20;
Matches 108; Conservative 71; Mismatches 156; Indels 63; Gaps 15;

QY 18 FG-CFVSGHDSATGELMIRNYFP-----PRAGEEGDETWGNDYEAR 57
DB 193 FGVGCHTAIDLKSGKVMKGYISIGPDQILVDEPKTSLGKPLADSSLKTEGD---Q 249
QY 58 WMTG---AMGQITDPTVNLVHGSTAVGPASETORGTGGTLYGTNTPRAVAPDGEIY 114
DB 250 WKTGGCTWGWFSYDPLKDLMTYVSG--NPSTNMPKQRPDGNW-SWTIARBPDIQMAK 306
QY 115 WRHQTLPDQWMDQCTEEMAVTNVDVOPSTEMEGLSINNAATGERRVLTGVPCKTGM 174
DB 307 WVYQMTPEHDEWDFGINEMILT-----OKPFG-----KDRPLTHFD-RNGFG 349
QY 175 WQFAETGEFLMARDTN-YQNMIESID-ENGIYTVNEDALIKELDEYD-----VCP 224
DB 350 YTLDRATGEVLVAEKFPVYVWMTATKVDLKGSTYGRPLVYSYSTEDQNEQVNSGICG 409
QY 225 TFLGGRDMPAALNPDGSIYFPLNNVCYDMAVDEFTSMDYNTSNVTKLP-PGK-DM 282
DB 410 AALGTAKQQAAPASPKGLFVYPLNHCMQDEPFRVYTPGQPVATLSMYPAPSSHG 469
QY 283 IGRIDAIDISTGRTLSVERAANYSPLVSTGGVLFNGGIDRYFRALSOETGELTMOTR 342
DB 470 MGNFIAMDNLQGIKIKSNPQFSAMGALATSGDVLFGYLEGFLKAVDSKTKGKELYRK 529
QY 343 LATVASGOAISYEVDGMOYVAIANG-----GVSYGSL 375
DB 530 TPGSIIQVMTYEBHKQHVAVLISGVGAGIGLAAGL 567

```

```

RESULT 4
Q9AQ48 PRELIMINARY; PRT; 629 AA.
AC Q9AQ48;
DT 01-JUN-2001 (Tremblrel, 17, Created)

```


RT fixing endosymbiont *Sinorhizobium meliloti*;
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AL603642; CAC48573.1; -
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424EFED15ADBE CRC64;

Query Match 18.4%; Score 374; DB 16; Length 601;
 Best Local Similarity 26.5%; Pred. No. 4.3e-20;
 Matches 113; Conservative 66; Mismatches 161; Indels 86; Gaps 15;

QY 2 VANGYVAGSTCOYSPFGCVSGHDSATGEELMRYFT-----PRAG 43
 DB 181 VKDKILVIGSGEFGVGRG-HYTAYSMDGKVLKRGYSKGPDSDLIDPEKTHLGKPYGK 239
 QY 44 EGGDEMGNDYEAHMTG---AMGQITDPTNLYHGSTAVGASSTOGRGTGTYNTR 100
 DB 240 DSGLTTEBD---QMKIGGGTTMGWYSTDPEENLYGTGPNSTWNPQR--PGDNRW-S 293
 QY 101 NTRFVRPDTEIWRHQTLPDMNDQCTFEMAVTNDVOPSTEMEGLQINPNATGE 160
 DB 294 MTFIARDVDTGMAKVLQMTPHDEWDYGVNEMILTEQIDGK-----D 337
 QY 161 RRVLLGVCKGTMMQFAETGEFL-----WADTNYQNKIE-----SI 199
 DB 338 RKLTHFD-RNGFGYMDRVGELLVAEKYDPTVMATEVMDPSKDYGRPYVAQYST 396
 QY 200 DENGIVTVEDALIKELDEYDCEPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVD 259
 DB 397 EQNG-----EDTNTT-----GVCPRALGTDDQAPATSPKTELFYPTNHVCMDYEPFR 445
 QY 260 QEFTSMDEVNTSNVTKLPFGKDM---IGRIDAIDISTGRTLSVERAANYSPVLSTGG 316
 DB 446 VSYTAGOPYVATISMYP-KDSHGKMGNFYAMDNKEGKIKWSPPEFSVWSGALATAGD 504
 QY 317 VLFNGTDRYFRALSOETGELMOTRLATVASGQAISEYDGMQYVALIAG-----GVSY 371
 DB 505 VFVGTLEGTAKAVDATGKELRYKTPSGVIGVMTYARAGKYAVLVSGVGMAGIGL 564
 QY 372 GSGINS 377
 DB 565 AAGLTN 570
 RESULT 7
 Q9F902 PRELIMINARY; PRT; 695 AA.
 AC Q9F902;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ACH.
 OS *Pseudomonas stutzeri* (*Pseudomonas perfectomarina*).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
 OC Pseudomonas.
 OX NCBI_Taxid=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT Identification and characterization of genes activated by 2-
 RT chloroethanol in *Pseudomonas stutzeri* BC-2;
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; -
 DR HSSP: Q92477; 1FLG.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR InterPro: IPR003045; CytC_heme_bind.
 DR InterPro: IPR003088; CytC.
 DR Pfam: PF01011; Bacterial_PQO; 6.
 DR Pfam: PF00034; cytochrome C; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.

SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 17.7%; Score 360.5; DB 2; Length 695;
 Best Local Similarity 27.9%; Pred. No. 5.6e-19;
 Matches 109; Conservative 57; Mismatches 173; Indels 51; Gaps 14;

QY 1 IVANGYVAGSTCOYSPFGCVSGHDSATGEELMRYFIPRAGE-----EGDET 49
 DB 185 VVKGKVIILNGAGAYVRG--FFSAYDAETGKMMARFYTP--GDPAPQYEPHELAAKAT 241
 QY 50 W-GNDYEAHMTG---AMGQITDPTNLYHGSTAVGASSTOGRGTGTYNTR 104
 DB 242 WKGDQY---WKLGGGTVWPMKAYDELDLLYGTGSGWNEIRSPGGDNLVLSIL 298
 QY 105 AVRPDTEIWRHQTLPDMNDQCTFEMAVTNDVOPSTEMEGLQINPNATGERVY 164
 DB 299 ALRPDSGKLLMHYQTPGETWDPATQQTTLATL-----ELDG-----KRRVL 342
 QY 165 TGVPCKGTMMQFAETGEFLARDINQNMISID--ENGIYVVEDALIKELDEYDV 222
 DB 343 MQAP-KNGFFYVLDRAITGELLISAKEGKYTMAEKVDLATGRPYEGRYERKEQVYVM-- 399
 QY 223 CPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVDQETSMDEVNT---SNVTKLP 278
 DB 400 -PSPFGAHNMHMSFNPQGLMITYIYQELIPGYRNAGATFKKIDGLNTGFSDTHELIP- 457
 QY 279 GKQMT-GRIDAIDISTGRTLSVERAANYSPVLSTGGVLFNGTDRYFRALSOETG 337
 DB 458 -RDVSGALAMPVPRQREARVPSPFWNGGTLSTAGNLVPGVADGGLAHYSADKQGR 516
 QY 338 LMQTRLATVASGQAISEYDGMQYVALIAG 367
 DB 517 LMSFAQTGIVAAPISFSLDGEQYAVVMAG 546

RESULT 8
 Q9L935 PRELIMINARY; PRT; 599 AA.
 AC Q9L935;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
 GN MXAF.
 OS *Methylovorus* sp. (strain SS1 / DSM 11726).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 OC Methylovorus.
 OX NCBI_Taxid=81683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SS1;
 RA Kim Y.M.;
 RT "Cloning and nucleotide sequence of mxaf gene of *Methylovorus* sp.
 RT strain SS1 DSM11726."
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF184915; AAD56237.2; -
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 SQ SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

Query Match 17.7%; Score 359; DB 2; Length 599;
 Best Local Similarity 25.9%; Pred. No. 5.9e-19;
 Matches 106; Conservative 70; Mismatches 161; Indels 72; Gaps 14;
 QY 10 GSTQGYSPF-----GC-----FVSGHDSATGEELMRYFT----- 39
 DB 177 GSTTQAFPAVKAYLVGCGSAGELGVGRGYATYADQKTGELVWRSFATGPDPNELAKDFN 236

QY 40 ---PRAGEE---GDETWDNDYEAARMTGA---WGQITVPVTLNLYHGSTAVGASSETORG 91
 DB 233 KNPFGGQMGSLGTKEWEE---GKIGGGTMMGWYAYPRLNLFYGGSGNAPWNETWR- 292
 QY 92 TPGGLYGTNTRFAVRPTGTETVWHQTLPRDNDQECTFEMATYNVVOGSTEHEGLOS 151
 DB 293 -PGDNKM-TMTIWARDVLTGAARKMGYOKTPIHDEMFGAVNDMLTDQVNGKTO----- 344
 QY 152 INPMNATGRRVLTGVPCGTGTMQFDEATGEFLVARDTN-YQNMIESIDENGITVNEED 210
 DB 345 -----PLTHVD-RNGMTIYLNKQTSYVQAAYDPAVNFVKYDLTGLPVPDP 393
 QY 211 AILKELDVE-YVCPTEFLGGRDPSAALNPDGSIYFIPLNVCYMAAVDEFTSMYDN 269
 DB 394 ESTRMDHGTNCPISAMFHNQGLDAYDPDSRTFFYGLNHICMDMEPFMLPYRAGQFV 453
 QY 270 TSNVTKLP----PGKDMIGRIDAIISTGRTLSYERAAANYSPLYSTGGVLFNGGTR 325
 DB 454 GATLAMYPPGNGPTKKEGQVLAQMDQVTEGKTKKTSVWGGLTATGGLVFNLTLDG 513
 QY 326 YFRALSGETGLMOTRLATVAGSQAISYVDGMQVYALAGGVSYSG 374
 DB 514 NITADTKNKKEIMTKKMSGALGAPMSYAYKKQYIA-----TNYGVG 557

RESULT 9
 ID 046444 PRELIMINARY; PRT; 708 AA.
 AC 046444;
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)
 DT 01-NOV-1998 (T-EMBLrel. 08, last sequence update)
 DT 01-DEC-2001 (T-EMBLrel. 19, last annotation update)
 DE QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 DE (EC 1.1.99.-) (OH-EDH).
 GN OHEDH.
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
 OX NCBI_TaxID=285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=96184549; PubMed=8654419;
 RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinochaemoprotein ethanol
 dehydrogenase of Comamonas testosteroni";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RP SEQUENCE OF 32-54 AND 477-490.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=95324580; PubMed=7601151;
 RA De Jong G.A.H., Geerlof A., Stoorvogel J., Jongejan J.A., De Vries S.,
 Duine J.A.;
 RT "Quinochaemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 Purification, characterization, and reconstitution of the apoenzyme
 with pyruiloquinoline quinone analogues";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=86242113; PubMed=3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;
 RT "Quinochaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 testosteroni";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION.
 RX MEDLINE=21536088; PubMed=11679760;
 RA Oudrie A., Hutzinger E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinochaemoprotein alcohol dehydrogenase from
 Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS

CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POQ, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CA57464.1; -
 DR HSSP: Q924J7; 1F1G.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR003086; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C.
 KW Signal; POQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match 17.5%; Score 355.5; DB 2; Length 708;
 Best Local Similarity 26.7%; Pred. No. 1.4e-18;
 Matches 105; Conservative 59; Mismatches 180; Indels 49; Gaps 12;

QY 6 VYVASTQYSPFGGFGVGHDSATGEELMRNFYF-----RAGEGDETWDNDYEA 56
 DB 208 VILGRGALEYVRG-VITAYDAETGERKRMFSYVGDSKPFEDSMKRAARTW--DPSG 264
 QY 57 RMTGA-----WGQITVPVTLNLYHGSTAVGASSETORGTPGTGLTYTNTRFAVRPTG 111
 DB 265 KMWEGGGGTMDSMTPFDALNTMYVTGNGSPMSHKYRSPGSGNLTASIALVADLPDG 324
 QY 112 EIVRHQTLPRPNWQECTFEMATYNVVOGSTEHEGLOSINPMNATGERVLTGVPCKT 171
 DB 325 KYKWHQETPGDNMDYTSQPMILADIKI-----AGKPRKVIILAP-KN 367
 QY 172 GTMMQFDEATGEFLVARDTNQNMIESIDENG-IYVNEADILKELDVEYDVCPTFLGR 230
 DB 368 GFEEVLDNRNGKFKISAKNVPVNMASGYDKHGKPIGI--AAARGSKFQDAVPPYGAH 424
 QY 231 DWPSAALNPDGSIYFIPLNVCYDMAAYDQ-BET-----SMDVYNSNTKL-PPGK 280
 DB 425 NMHPMSFNQDTGLVLYPAQNVPNVLDKDKWENQAGPCKPOSQGMNTAKFFNNEPPKS 484
 QY 281 DWIGRIDAIISTGRTLSYERAAANYSPLYSTGGVLFNGGTRDYFRALSGETGLTMO 340
 DB 485 KPPGRLAMPVQAOKAMSEVHSPNNGGTLTLAGNVFQGTADGRVAVYHNAATGEKIME 544
 QY 341 TRLATVAGSQAISYVDGMQVYALAGGVSYG 372
 DB 545 APGTGVAAAPSTYVMDGRQYSAVAGWGGVYG 577
 RESULT 10
 ID 024759 PRELIMINARY; PRT; 633 AA.
 AC 024759;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, last sequence update)
 DT 01-DEC-2001 (T-EMBLrel. 19, last annotation update)
 DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
 GN KJAF.
 OS Hyphomicrobium methylovorum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Hyphomicrobium group; Hyphomicrobium.
 OX NCBI_TaxID=84;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-GM2;
 MEDLINE-97437202; PubMed-9311140;
 RA Tanaka Y., Yoshida T., Matanabe K., Izumi Y., Mitsunaga T.;
 RT "Cloning and analysis of methanol oxidation genes in the methylotroph
 RT *Hyphomicrobium methylovorum* GM2.";
 RL FEMS Microbiol. Lett. 154:397-401(1997).
 DR EMBL: AB004097; BAA23272.1; -.
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 KW Signal; Oxidoreductase.
 FT SIGNAL 1 34
 FT CHAIN 35 633
 FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;

Query Match 17.2%; Score 350.5; DB 2; Length 633;
 Best Local Similarity 26.2%; Pred. No. 2.8e-18;
 Matches 111; Conservative 70; Mismatches 169; Indels 73; Gaps 16;

QY 2 VANGVIAGSTQYSPFGC-----VSGHDSA-----TGEELMKNYFTPRAGEGD----- 39
 DB 181 VENGDIKVGQILQAPVYVHDALIVSGSAGELGVAGHVAIVNKTGEQAMRYVATGPDEE 240
 QY 40 -----PRAGEE--GDETMGNDYEAARMTGA--WGQITYPDVTNLVHYGSTAVG 83
 DB 241 IGLADDFNSANPHYGKGLGATWEGD---AMKIGGGTNMGVAYADPQANLIYSSGSPA 297
 QY 84 PASETORGTPGGLTGTWTRFAVRPDGEIYMRHQTLPRDNDDECTEEMATVNDVOPS 143
 DB 298 PWNEMMR--PGDNKW-TTITARADYGEKMKFGYOKTPHDEMDAGVNVIMLSE-----Q 349
 QY 144 TEMBGLSINPNAATGERRVLTGVPCKTGMOPDAETGEFLMA---RDTNYOMIESID 200
 DB 350 TDKE-----GKKRKLTLHPDRNGIYITLDRNGDISDKDDT--VAVFKAHD 396
 QY 201 ENGIYTVNEDALIKELVYE-YDVCEPTFLGGKMPSAALNDPSGIYFIPLNNVCYDMAVD 259
 DB 397 LKSGLPVRDPEFGTRMDHKGTETLCPSAMGYHNOGHSYDPTKOLFEMGINHICMDWPFM 456
 QY 260 QEFTSMDEVYNTSVTKLP-PGKDM-----IGRIDAIDISGRILMSVERAANYSPVLS 313
 DB 457 LPRAGQEFVCAITLMTPGKGRGQNLGLGQIKAYNAITNTYKWEHMERFSVWGTLAT 516
 QY 314 GGVLFNGGIDRFRALSOETGETLMQTRLATVASGOAISYEVDGMQVYIAGG-GVSYG 372
 DB 517 AGNLVFEYGTLDGLKANSDDTGLIMKHLPSGVITGYPMTEYHKGYOYIAVMSGVGWP 576
 QY 373 SGL 375
 DB 577 VGL 579

RESULT 11
 ID 053362 PRELIMINARY; PRT; 742 AA.
 AC 053362; 044159;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
 OS Acetobacter pasteurianus (Acetobacter turbidans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 NCBI_TaxID-438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NC11380;
 RX MEDLINE-94042848; PubMed-8226628;
 Takemura H., Kondo K., Horinouchi S., Beppu T.;

RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
 RT pasteurianus.";
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL: D13893; BAA40252.1; -.
 DR HSSP: Q92407; 1FG.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_6.
 DR PROSITE: PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 KW SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 16.6%; Score 338; DB 2; Length 742;
 Best Local Similarity 26.7%; Pred. No. 3.1e-17;
 Matches 114; Conservative 62; Mismatches 165; Indels 86; Gaps 16;

QY 2 VANGVIAGSTQYSPFGC--FVSGHDSATGEELMKNYFTPRAGEGD----- 47
 DB 203 VAKGLVILGN--GSGEARGFVSAPFAETGKLMKRYTPNNKNEPDHADVADNVLMSKA 260
 QY 48 -ETWGNDEYARMT-----GAMQITYPDVTNLVHYGSTAVGPASETORGTPGGLT--- 98
 DB 261 YKTWGP--KGAWVROGGGGTVMDSLYVDPVSDLIY--LAVG-----NGSPNNYRSE 309
 QY 99 --GTN---TRFAVRPDGEIYMRHQTLPRDNDDECTEEMATVNDVOPSTEMBGLSI 152
 DB 310 GIGSNPLFGLSVLAKPEGTGEYVNHFOATPDQDDYNSVOOIMTLDPV----- 357
 QY 153 NPNAAATGE--RRVLTGVPCKTGMOPDAETGEFLMARDTNYOMIESIDE-NGIYTVNED 210
 DB 358 -----NGEMRHVYIWHAP-KNGFEYVLDATGEELACKNIVYQWANGLDPLTGRPIYND 411
 QY 211 AIILKELDEVDCEPTFLGGKMPSAALNDPSGIYFIPLNNVCY-----DMA 257
 DB 412 GLTYLTGKFTWYGLPGLGANEFMGNAVSPKTHLVYPAHQIPFGKNQVGGKPPDAMN 471
 QY 258 VDEFTSMDEVYNTSVTKLP-PGKDM-----IGRIDAIDISGRILMSVERAANYSPVLS 317
 DB 472 VGLDMKKNLPDTPPE-ARAYIKDLHGLLLANDPVKMETWYKIDIKGPMNGVLVATGDL 530
 QY 318 LKNGGIDRFRALSOETGETLMQTRLATVASGOAISYEVDGMQVYIAGG-GVSYG 367
 DB 531 LFQGLANGEFHAYDATNGSDLYKFDQSGIITASPMYISVNGQYVAVEVWGNGIPI 590
 QY 368 GVSYSG 374
 DB 591 GVGRISG 597

RESULT 12
 ID 059540 PRELIMINARY; PRT; 573 AA.
 AC 059540;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE HEAVY SUBUNIT (EC 1.1.99.8).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 OC Methylophilus.
 NCBI_TaxID-17;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W3A1;
 RA Xia Z.-X., Dai W.-W., Zhang Y.-F., He Y.-N., White S.A., Boyd G.D.,
 Mathews F.S.;
 RT "Determination of the Gene Sequence and the Three-dimensional
 RT Structure at 2.4 Å resolution of Methanol Dehydrogenase from
 RT *Methylophilus methylotrophus*.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U41040; AAA83765.1; -.
 DR HSSP: P38539; 4AAH.

Db 341 ----RKVILHAP-KNGFEFVIDRTDGEKFIISAQNFVFNWATGDENGRPIENEGAMPGH 395
 QY 216 LDVEYDVCPTFLGGDMPSAALNPDSCITFIPIANVCYMAVDOFTSMYNTSN--- 272
 Db 396 LSRF---PAPSAKIMHMSYSPOGLAFPAQNTPL-VLOEDKNWS---YNQAPQO 447
 QY 273 -----VTKLPKGMIGRIDAIDISTGRTMSVERAAANYSPLVSTGGVLFN 320
 Db 448 AMAGIGMNLGMLVNPBPASQPPGRILAMPVOQKEVWKKEHVSFNGGTLVYAGNVYHQ 507
 QY 321 GGDTRFRLASOETGETLMQTRLATYASGOAISYEVDGMOYVAILA-GGVSYSYS 373
 Db 508 GTADARLLAFDARDGKELMSAPMGTVIAPVYEVGDKQYVSIANGMGVYGN 561

RESULT 15

09AGW3 PRELIMINARY; PRT; 623 AA.
 AC 09AGW3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
 RT "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF326086; AAK15506.1; .
 DR HSSP; Q924J7; 1FLG.
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ; 4.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 13.9%; Score 281.5; DB 2; Length 623;
 Best Local Similarity 24.2%; Pred. No. 4.6e-13;

Matches 105; Conservative 65; Mismatches 161; Indels 103; Gaps 18;

QY 6 VIVAGSTCOYSPFC--FVSGHDSATGEELMRNYFI-----PAGEEGDEITNG 51
 Db 199 LTVHGG--GDEFGVGMLFARDPDTGEVYARPMVEGHGRLNGKDSIPTGDKAPSWP 256
 QY 52 NDYEA-----RWMTGA--WGQITYPVTNLVHYGSTAVGPASETQRTGP-----GT 96
 Db 257 DDPSPTGKVEAMSGGGAPWQTFASFDVENNMVYIGANPAPWMTKRTAPGDDPRMWS 316
 QY 97 LYGTNTRFAVAPDTGEIYVWRHOTLPKRDNDQECTFEEMVTN-----VDYQSTEME 147
 Db 317 LF-TSGQAYVDASTGELKGYQHTPNDAMPSCGNNSVLFERYKDPKTKMNASAHADR 375
 QY 148 GLOSIN-----PNAATGERVLTGVPCKTGTMMQ--FDAETGEFLMARDNTYQN 194
 Db 376 GFEFVTRDRMLAKGAGYPNKPST--LIGAMPVVDGITWASGFDLTKGPI-EKDNRPQ 431
 QY 195 MIESIDENGIVTNEDEILKELDEYDVCTFLGGDMPSAALNPDSCITFIPIANVCYD 254
 Db 432 PKGADGESIFVS-----PPFLGNTWHPMSYSPTGLFYIPANHAMD 476
 QY 255 MAVDOFTSMYNTSNVT-----KLPPG-----KDMIGRIDAIDISTGRTMSVE 301
 Db 477 -----YMTENTYTKAGSAYLGGFRIKNLFDHVGILRAIDPSPARSLSGAOG 523
 QY 302 RAANTSPVLSTGGVLFNGGTDYFRALSOETGETLMQTRLATYASGOAISYEVDGMOY 361

Db 524 RVPA-VAGTLTAGGWFTGTSDSGLAKAFDAKNGKELMKFOTGSGVVSVPVTWEMDEQY 582
 QY 362 VAIAGGVSYSGSL 375
 Db 583 VAIQSG---YGVAV 593

Search completed: May 24, 2002, 10:17:33
 Job time: 746 sec

Fri May 24 11:27:12 2002

us-08-934-506a-5_copy_180_556.rsp

